6-16-2011

Estimating Anthropometric Marker Locations from 3-D LADAR Point Clouds

Matthew J. Maier

ESTIMATING ANTHROPOMETRIC MARKER LOCATIONS FROM 3-D LADAR POINT CLOUDS

THESIS

Matthew J. Maier, Captain, USAF

AFIT/GE/ENG/11-27

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GE/ENG/11-27

ESTIMATING ANTHROPOMETRIC MARKER LOCATIONS FROM 3-D LADAR POINT CLOUDS

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Matthew J. Maier, BSEE

Captain, USAF

June 2011

AFIT/GE/ENG/11-27

ESTIMATING ANTHROPOMETRIC MARKER LOCATIONS FROM 3-D LADAR POINT CLOUDS

Matthew J. Maier, BSEE
Captain, USAF

Approved:

Maj Michael J. Mendenhall, PhD (Chairman)	9-JUNE-2011 Date
Dr. Richard K. Martin (Member)	09 Jun 2011 Date
Dr. Gilbert L. Peterson (Member)	9 JUNE 2011 Date

Abstract

With a greater emphasis being placed on rapidly identifying non-conventional adversaries, the U.S. armed forces have placed a high priority on expanding and developing dismount detection capabilities. Current work in this area utilizes a multispectral imaging system to exploit spectral properties in order to detect areas of exposed skin and identify clothing properties.

An area of interest for improving the characterization portion of the system is in extracting anthropometric markers from a Laser Detection and Ranging (LADAR) point cloud. Analyzing anthropometric markers is a common means of studying how a human moves and has been shown to provide good results in determining certain demographic information about the subject. This research examines a marker extraction method utilizing principal component analysis (PCA), self-organizing maps (SOM), alpha hulls, and basic anthropometric knowledge. The data set is comprised of 3-D body scans and truth marker locations from the Civilian American and European Surface Anthropometry Resource (CAESAR) Project. The performance of the extraction algorithm is tested by performing gender classification with the calculated markers. Classification is performed using linear discriminant analysis. The measured classification accuracy using the estimated marker locations is approximately 78% whereas the classification accuracy using the truth marker locations is 88%. This indicates that the marker estimation algorithm developed in this research performs well and shows that fusing LADAR data with the multispectral detection system's data is a promising path for improving the dismount detection and identification problem.

Acknowledgements

I would like to thank all those working at the Air Force Institute of Technology. Without all of the support staff performing their jobs, the challenging task of doing research would be that much harder. I would also like to thank all of the professors with whom I had the benefit of having a class. Their knowledge, dedication, and passion for the subjects made the classes more instructive and enjoyable.

Thanks go to my thesis committee: Dr. Gilbert Peterson and Dr. Richard Martin, whose classes were challenging but very informative, and who took an interest in reviewing and approving my research. Special thanks goes to my academic advisor, Maj Michael Mendenhall. His guidance throughout this process is much appreciated and was indispensible to achieving this goal.

Thanks also goes to the other students that were a part of the various study groups throughout the courses and especially to the other members of the research group, Jonathon Climer and Bradley Koch, whose support was invaluable. A final thanks goes to my sister who was crucial to allowing me to focus on my goal.

Matthew J. Maier

Table of Contents

Page

List of Figures

Figure Page

Figure Page

List of Tables

ESTIMATING ANTHROPOMETRIC MARKER LOCATIONS FROM 3-D LADAR POINT CLOUDS

1. Introduction

The United States Air Force Science & Technology Program continues to stress the importance of research into persistent surveillance capabilities that provide the warfighter with all-weather, around-the-clock, continuous tracking of dismounts. Dismount tracking is useful for inferring intent which can provide information about whether the dismount is exhibiting suspicious behavior [40].

Significant advancements in the area of dismount detection have been achieved by the Sensors Exploitation Research Group at the Air Force Institute of Technology [8, 14, 15, 45, 65, 67]. The work in [8, 15, 45, 65, 67] has focused on skin-cued dismount detection through the use of multispectral analysis. An example of the results from the skin-cued dismount detector is shown in Fig. 1. Another element in the overall dismount detection system is clothing detection and characterization [14].

An opportunity for new research of benefit to solve the dismount detection problem is the fusing of Laser Detection and Ranging (LADAR) data with the existing approaches to provide anthropometric information that can aid the characterization piece of the dismount detection system. The limited nature of the literature available on applying marker estimation techniques to LADAR data indicates the novelty of the concept of anthropometric marker extraction.

1.1 Problem Statement

The goal of this thesis is to advance the state-of-the-art dismount detection and characterization system. Focusing on the characterization component through the segmentation of 3-D LADAR point cloud data to build a human abstraction based on estimated anthro-

Figure 1. An example of the results from the skin-cued dismount detection system with a bounding box placed around the detected dismount.

pometric markers. Two methods of verification are employed in evaluating the performance of the marker estimation algorithm. The first is through the measured offset between the estimated and the actual marker location. The second method is a verification of the utility of the marker estimates through the use of a gender classification algorithm.

1.2 Scope

The scope of this thesis effort must be limited in order to accomplish the previously stated goals. Accordingly, the scope of the research centers on the following:

- Obtain anthropometric marker locations from 3-D LADAR data;
- Limit the classification process to that of gender; and
- Analyze the performance of the marker estimation algorithm through classification accuracy and error measurements.

The Civilian American and European Surface Anthropometry Resource (CAESAR) Project is the source for the 3-D scans [72]. Associated with this research, there are several challenges pertaining to the data. The first is that the available data set consists of only 3-D

scans of adults. Secondly, the scans are not racially diverse. The data set is comprised of only populations from North America, Italy, and the Netherlands. This limits the possibility of performing classification based on race/nationality characteristics. The demographics for the database are provided in Section 4.1.

The marker estimation alogrithm is based on a conceptual framework presented in Chapter 3. The 3-D point cloud data is projected into a dimensionally reduced space and normalized. These steps are detailed in Section 3.2. Section 3.3.1 describes the Self-Organizing Map that is applied to the projected data and extracts cluster centers from the distribution of the projected data. The calculated cluster centers are then refined based on the projected data's geometry which is discussed in Section 3.3.2. Further refinement of the cluster centers is performed using skeletal information, covered in Section 3.3.3, and anthropometric information, discussed in Section 3.3.4.

1.3 Document Organization

This thesis consists of five chapters. This first chapter motivates the research effort, states the problem, and presents the scope and limitations associated with the research.

The second chapter provides a discussion of literature related to problem of segmenting 3-D LADAR point cloud data. The chapter also contains the background information necessary to understand the structure and components of the developed marker extraction algorithm.

The third chapter presents the methodology by describing the marker estimation algorithm from the 3-D LADAR data through the various stages of processing to the resulting anthropometric marker estimates. This chapter specifically focuses on the algorithm's primary components: Principal Component Analysis (PCA), Self-Organizing Maps (SOM), Alpha Hulls, and the process of marker refinement.

The fourth chapter provides the details of the parameters specific to each stage of the marker estimation algorithm. The performance of the algorithm based on gender classification is presented in this chapter as well.

The fifth chapter summarizes the results, discusses the impact of this thesis' contribution

to current dismount detection and characterization efforts, and provides recommendations for areas of future work in improving the marker estimation algorithm.

2. Background

This chapter provides an overview of the problem of segmenting 3-D point clouds and how others have approached the problem. Additionally, this chapter provides the necessary background information on the pattern recognition, neural network, and classification tools implemented in this thesis.

The chapter begins with a brief overview of the notation and terminology conventions used in this thesis. This is followed by an overview of the nature of point cloud data and a review of current approaches to the segmentation problem. The next section deals with the tools used in this thesis to build the segmentation algorithm. The final portion of this chapter discusses previous related work and the issues with classification.

2.1 The Nature of LADAR Point Cloud Data

Light Detection and Ranging (LiDAR) systems, of which LADAR systems are a subset, are a common tool used to collect ranging information used in a myriad of fields, including terrestrial mapping and target tracking. LiDAR, LADAR, and RADAR are all ranging systems that transmit and receive electromagnetic information. In the case of LADAR, the transmitted energy falls into the frequency band associated with laser energy. A simplified view of how LADAR performs is presented in Fig. 2. It shows that a pulse is transmitted and an approximately co-located receiver detects the returned pulse. The time delay between the transmission and reception of the pulse determines the distance to the scanned object. This 1-D example can then be expanded to a 2-D array of receivers that can provide a 3-D representation of the scanned object. The output of the LADAR system used in the human body scans is a dense point cloud representing the location of the thresholded signal return.

Typical LiDAR segmentation can be classified into two categories: part-type segmentation and patch-type segmentation. Part-type segmentation is commonly used in terrestrial mapping to extract basic geometric shapes from the data [71]. Patch-type methods strive to segment the data into homogeneous regions based on similarities between local points or the proximity of the points [88]. The segmentation of the human body scans falls into the

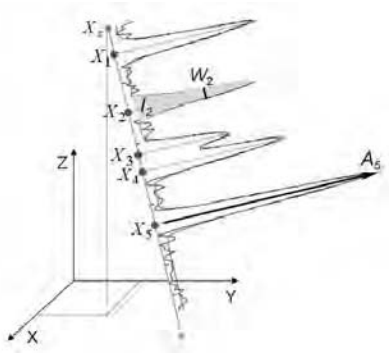


Figure 2. An example of the use of a LiDAR system to map terrain. The LiDAR transmits a light pulse which is reflected back to the receiver by the various surfaces in the path of the light pulse corresponding to the spikes in the waveform.

latter category of segmentation.

3-D LiDAR segmentation has interest from a number of diverse fields. The geo-spatial and remote sensing community often look to segment data for terrestrial mapping applications [12, 61, 88]. The computer vision community uses segmentation in the area of vision-aided navigation [38]. Pertaining to the segmentation of LiDAR data from human body scans, there is significant interest from the perspective of the computer graphics field in order to develop more realistic computer models of humans for the entertainment and simulation industries [9, 31, 50]. Additionally, the desire to improve automated human identification and recognition has provided more avenues for further progress in the area of point cloud segmentation.

Current strategies that address the problem of segmenting 3-D point cloud data can be generalized into two basic methods: segmentation from the point cloud directly or surfacing the point cloud before segmentation. The former approach benefits from a reduction in the amount of algorithm preprocessing necessary to conduct segmentation, whereas the latter approach attempts to take advantage of the underlying curved nature and geometric properties of the object to aid in identification and segmentation [2]. With the end goal of extracting anthropometric markers from point cloud data, a part of this research examines the recent research into LiDAR segmentation.

2.2 Segmentation Approaches

Much of the published literature pertaining to the segmentation of 3-D point cloud data falls into two basic categories: segmenting point cloud data directly or segmenting a surfaced representation of the point cloud. Often, the decision of which method to use is based on the academic field using the information. Geo-spatial and vision-based navigation applications tend to utilize raw point cloud data processing methods whereas the computer graphics community focuses on surfacing based methods. One reason for this difference is due to the format of the data returned from the sensor. Typically, fielded LiDAR sensors for geo-sensing applications return either waveforms or point clouds; however, many 3-D scanners employed in computer graphics applications process the data and generate polygonally meshed surfaces. Whether it is with the intent of incorporating a surfacing algorithm into geo-sensing applications or an effort to improve the surfacing algorithms designed into the 3-D scanner software, there is significant research in the area of surfacing point cloud data [28, 47, 60, 66].

2.2.1 Segmentation from Point Clouds.

Based on the previously described distinctions, the literature pertaining to implementing clustering algorithms directly on point clouds from the remote sensing community generally addresses terrestrial mapping problems. A large portion of the research in this area is split into papers dealing with improvements or variations to the K-means algorithm [11, 12] with others focusing on research into implementing or adapting Iterative Self-Organizing Data Analysis Techniques (ISODATA) algorithms [61].

Direct segmentation from point cloud data is also addressed from the computer vision community, generally from the perspective of vision-aided navigation for robots but also in the area of human tracking and motion capture [13, 29]. Often, these methods focus on K-Nearest Neighbor (KNN) based approaches [44], such as Clustering Large Applications based on RANdomized Search (CLARANS) [64] and Density-Based Spatial Clustering of Applications with Noise (DBSCANS) algorithms [21].

A direct application of processing point cloud data of human body scans is addressed in [16] and builds upon the capability of spectral embeddings to map 3-D shapes into lower dimensional manifolds [3, 82]. Specifically, the author explores the capability of locally linear embedding (LLE) to reveal the intrinsic geometry of the articulated shape. Additionally, the author compares the performance of LLE to Isomap, a low-dimensional embedding method in which geodesic distances on a weighted graph are incorporated with metric multidimensional scaling [36].

2.2.2 Segmentation From Surfaces.

When approaching segmentation from the perspective of surfacing, the published work can be divided into at least three categories: graph theory based methods, extensions of image processing techniques, and perceptual segmentation. In general, approaches considering segmentation from a surface assume that the surface was provided as the output from the scanner or accomplished as a preprocessing step with the nature of the surfacing algorithm left as a separate area of research.

A popular graph theory method for segmenting a surfaced point cloud is the Reeb Graph [93]. The Reeb Graph, a mathematical construct based on Morse theory, represents the topological skeleton of a multi-dimensional object. A height function is used to calculate level-set curves which represent a vertex in a graph, producing a skeleton of the object [19]. Implementations of a basic Reeb Graph are generally sensitive to noise in the data. As such, the numerous adaptations of this approach include the following: examining various height functions [81], using a discrete Reeb Graph [86, 92], and affine invariant Reeb graphs [62].

A second approach to segmenting a surface extends 2-D image processing techniques to 3-D applications. This approach includes using the watershed algorithm for segmentation [52], extending binary morphologic transforms and filters [73], and applying the Medial Axis Transformation (MAT) to the problem [80]. The first approach involves examining the local curvature minima in order to derive the segmentation criterion. The latter two methods produce a skeleton which can be segmented through graph-based characterization methods

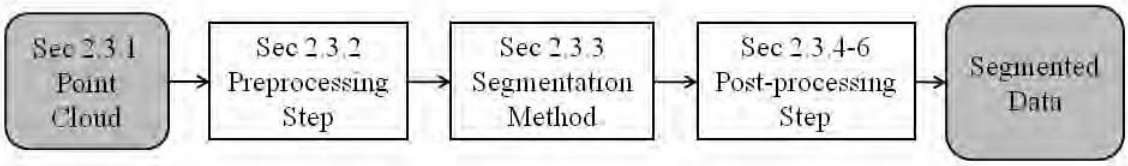


Figure 3. Fundamental diagram of the overall segmentation problem for the two approaches: segmenting from a point cloud and from a surface.

similar to those used in Reeb Graph techniques.

The third approach draws on similarities to the way in which human perception segments objects [48, 91]. This approach is based on minima theory and how human perception decomposes an image. From the perspective of the literature, this appears to have been a popular approach in the initial stages of addressing the problem, but there has been little research in this direction in recent years.

2.3 Conceptual Paradigm for the Marker Extraction Problem

In approaching the problem of extracting anthropometric markers directly from point cloud data, the problem can be broken down into several fundamental processes. A representation of the approach is depicted in Fig. 3.

2.3.1 Point Cloud Data Analysis.

In exploring several of the algorithms developed from the background research, the 3-D point cloud data used in testing the developed algorithms is from the Civilian American and European Surface Anthropometry Resource (CAESAR) Project [72]. Each data set is large, on the order of hundreds of thousands of points. An example of the point cloud data is shown in Fig. 4. The data was randomly downsampled for the purpose of display. There are several challenges in dealing with the collected data. The scans were conducted in three separate locations: North America, Italy, and the Netherlands, using different scanners and different data collection teams, both resulting in inconsistency in the values for marker locations. Additionally, the fidelity of the scans varies with some scans having sparse data points in regions of the body or missing body parts entirely.

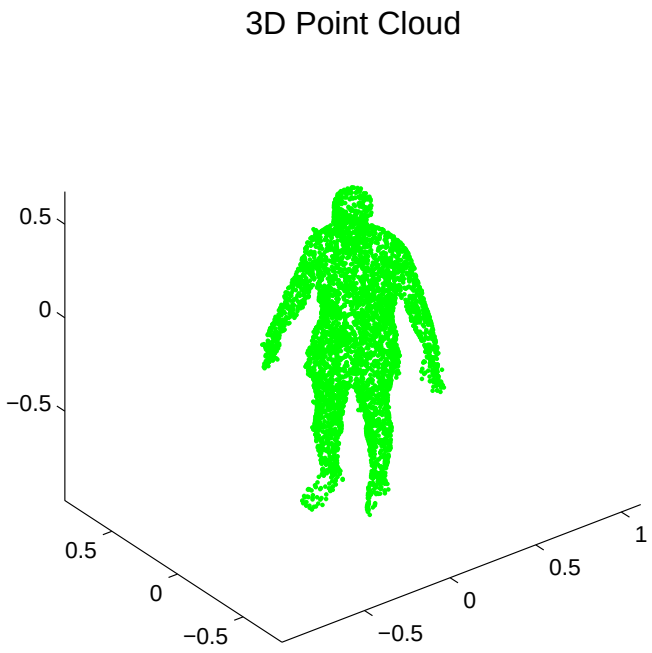


Figure 4. Example 3-D point cloud from the AFRL/RWPA Civilian American and European Surface Anthropometry Resource (CAESAR) Project database [72].

2.3.2 Projection Methods.

As part of the preprocessing phase, dimensionality reduction is often used with the goal of simplifying the complexity of the problem. There are a number of dimensionality reducing, or projection methods, used in the field of pattern recognition [24, 77] including Sammon's nonlinear mapping, multidimensional scaling, independent component analysis, and even self-organizing maps. However, the focus from the background research is on principal component analysis and locally linear embedding, discussed in the following sections.

2.3.2.1 Principal Component Analysis.

Generally, principal component analysis (PCA) is employed in dimensionality reduction, data compression, and feature extraction problems. However, it can be useful in data analysis/visualization efforts. Applying PCA to the 3-D point cloud data can provide information about various angles of rotation and the eigenvectors derived from the covariance of the point cloud provides the principal axes of the data. In general, PCA is implemented

using the following equation:

$$\mathbf{Y} = \mathbf{\Phi}^H \mathbf{X}, \tag{1}$$

where Φ is the eigenvector matrix found from the covariance matrix of $\mathbf{X}$, $\mathbf{X}$ is the original data, and $\mathbf{Y}$ is the output data in the transformed space. In this case, the desired result is the eigenvectors from the decomposition of the covariance matrix. After projecting the 3-D data onto the first and second principal components, the eigenvectors of the autocorrelation matrix provide the two vectors that describe the major and minor axis of the data distribution.

Examining the data in the transform space using all three principal components, as shown in Fig. 5, does not provide much in the way of additional insight into the data. Using the first two principal components produces the results in Fig. 6. Since the principal components are ordered in descending order of variance, the first component corresponds to the subject's height, the second component to the subject's breadth, and the third component to the subject's depth. The plane formed by the first and second principal components corresponds to the coronal plane as shown in Fig. 7. The coronal plane is likely sufficient for determining marker locations [26].

2.3.2.2 Locally Linear Embedding.

An interesting non-linear approach to the problem of segmenting a 3-D point cloud of a human subject is discussed in [16], in which locally linear embedding (LLE) is used. The rationale behind this approach is that, in general, multidimensional scaling (MDS) and PCA both fail to provide a method to reduce the dimensionality of non-linear point cloud data because they model linear variabilities over the entire data set. Due to the inherent nonlinearity of the human body scans, utilizing a dimensionality reduction tool designed to operate on nonlinear manifolds makes sense. LLE is such a tool. It is an unsupervised learning algorithm that "computes low dimensional, neighborhood preserving embeddings of high dimensional data" [74].

Local Linear Embedding Description

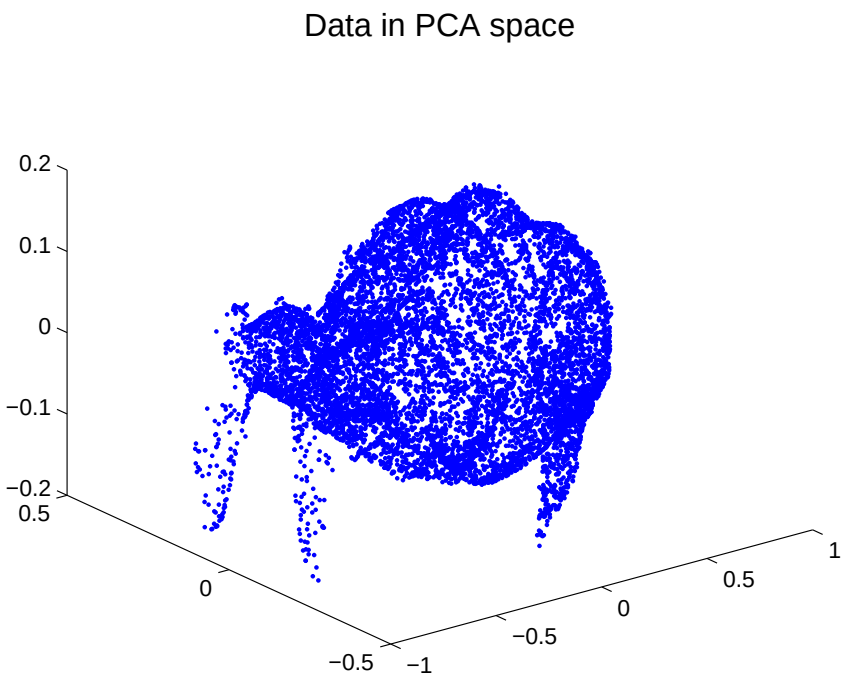


Figure 5. The results of transforming the data into PCA space while retaining all three principal components. The figure's rotation changes since the axes of the body align with the principal component axes.

LLE makes use of the approximate linear relationship of a point set of nearest neighbors, assigning a set of weights to each point based on its linear relationship with its neighbors. An eigenvector optimization algorithm is used to find a lower dimensional embedding of the points, preserving the linear relationship of a point with its nearest neighbors [74]. The results of this transform can be seen in Fig. 8. As can be seen, the LLE transform improves upon the PCA transform in terms of providing more definitively segmentable regions. From Fig. 9 we can see that LLE also provides a means of reducing the 3-D point cloud dimensionality while maintaining a means to segment the data into natural groupings.

Evaluation of Local Linear Embedding

A big discriminator in evaluating the success of the various segmentation algorithms lies in their computational efficiency. The computational intensity of KNN determination and eigenvector decomposition results in the LLE algorithm being very slow. In order to achieve results, the data set was randomly downsampled, retaining only 2.5% of the original data set.

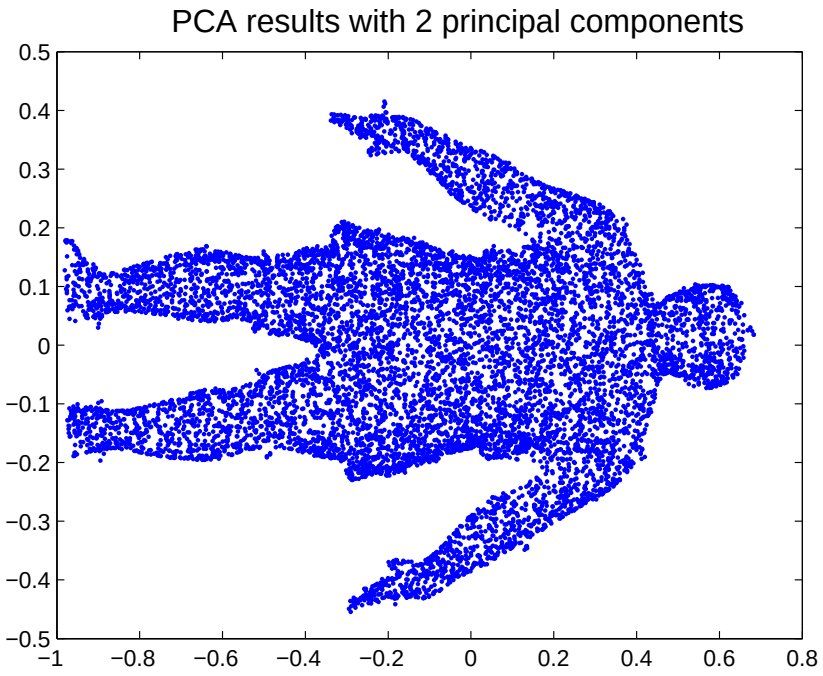


Figure 6. The transformed data using the first two principal components in the projection of the 3-D LADAR point cloud data.

LLE provides a good method for segmenting the data as well as a way to provide dimensionality reduction. Using LLE, in conjunction with a point representation method such as the mean, may provide suitable results when extracting marker locations from the point cloud data. However, the computational intensity of the KNN portion of the algorithm is a drawback. Due to the KNN approach of most of the vision-navigation-based clustering algorithms, they suffer from similar problems when applied to the large human body scan data.

2.3.3 Clustering Methods.

Moving into the area of segmentation, there are a number of hierarchical and agglomerative clustering methods available. However, as part of this research the focus is on examining the K-means, formal element, and self-organizing map algorithms.

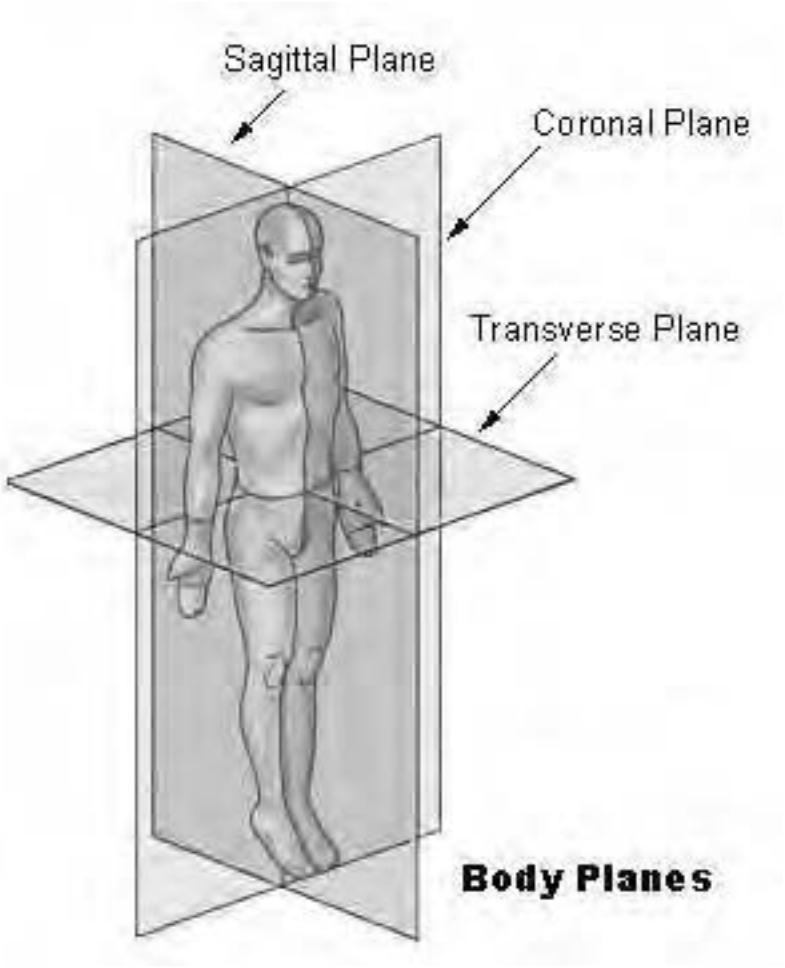


Figure 7. The nomenclature and orientation of the primary planes associated with describing planar locations of anatomical features and views.

2.3.3.1 K-Means.

K-Means Description

In the K-means approach, K randomly generated points are chosen and the algorithm iteratively updates the location of the K points to correspond to the density centers of the updated cluster regions. One of the problems with this method when applied to the human body scan data is that the natural anthropometric information is not taken into account when clustering the data. This may result in instances where a mean of a region falls outside the contained volume of the scanned human body.

Implementing an unsupervised K-means algorithm with $K = 20$ produces the results

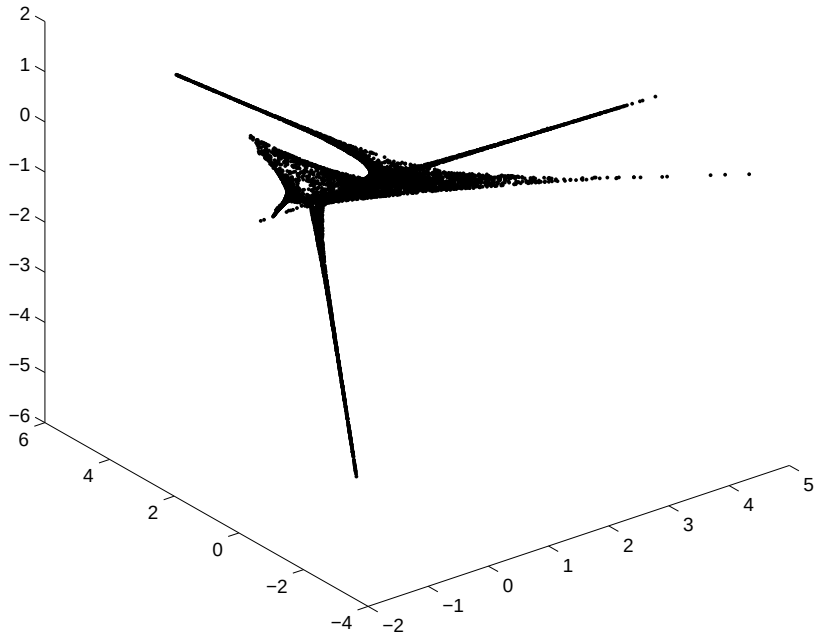


Figure 8. The 3-D point cloud data in the 3-D LLE domain.

shown in Fig. 10. The algorithm was run using the entire data set but was randomly downsampled in the plot to show where the means align within the data. As shown in Fig. 10, there are two means that fall outside the volume of the human body.

Evaluation of K-Means

Due to the unsupervised nature of the K-Means algorithm, its ability to extract marker locations from the data comes into question. However, if some anthropometric *a priori* knowledge is used in conjunction with it, the K-Means algorithm could provide suitable results.

2.3.3.2 Family of FOREL Clustering Algorithms.

Background

The formal element (FOREL) algorithm [69] is a taxonomy-based algorithm that falls

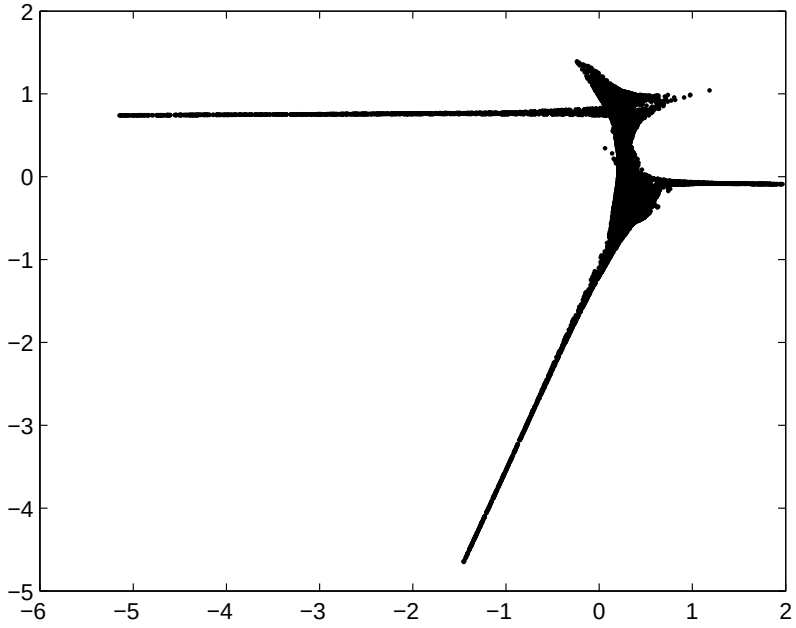


Figure 9. The 3-D point cloud data projected into the 2-D LLE domain.

into the category of agglomerative clustering [32]. It was first used in statistical data processing of palaeontology data [32]. In recent years, it has seen increased use in data mining efforts in the medical field, specifically in genetic and cancer research [68].

The goal of the FOREL algorithm is to divide m objects into K taxons based on the criterion of cluster compactness, assuming that $K < m$. Each taxon is comprised of a cluster of relatively close data points and represented by a central data point. The other taxons are comprised of data points that are maximally separated from the other taxons. By growing the radius of the hypersphere from a minimal value to a maximum value corresponding to half of the maximal distance, the data can be explored and the decreasing number of taxons tracked.

Though there are several variations of the basic algorithm [49, 94, 95], the general scheme of a FOREL algorithm is as follows:

1. For a fixed threshold radius, T, place the cluster center at a random data point.

2. Determine which data points lie within the sphere defined by the radius, T.

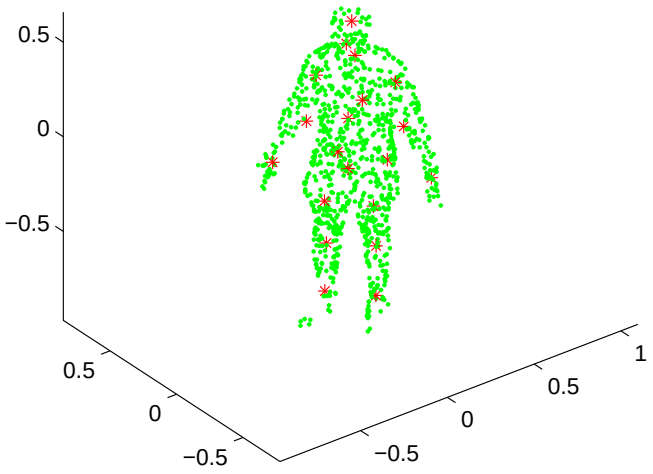


Figure 10. The results from the K-means algorithm. Notice that the two means near the subject's armpits fall outside the region of the body.

3. Using all data points that fall within the sphere, calculate the center of gravity.

4. Move the center to this new location, determine the new set of internal points and calculate the new center.

5. Repeat steps 1-4 until the center of gravity ceases varying. The sphere is referred to as a taxon, and all of the enclosed data points are represented by the calculated center of gravity.

6. Perform this process for the rest of the input data points until all data points are included in a taxon.

Litinskii-Romanov Algorithm

In [49], the authors propose a variation of the FOREL algorithm in which only the between data point distances are used in a neural network algorithm to extract class centers.

An N x N matrix consisting of the Euclidean distances between the m-dimensional points is calculated. At the beginning of the algorithm, each data point is represented by a

neuron with initial activity, $S_i(0)$. The weight matrix is comprised of connections between the ith and jth neurons, w_{ij}, given by the following:

$$w_{ij} = \begin{cases} \frac{T^2}{D_{ij}^2+T^2} & \text{when } D_{ij} \leq T, \\ 0 & \text{when } D_{ij} > T. \end{cases} \tag{2}$$

Eqn. 2 states that if the distance between points is greater than T then there are no connections between the neurons. However, if the distance between points is less than the threshold radius, T, the weight for the associated neurons is defined as the ratio of the squared threshold radius to the sum of the squared inter-neuron distance and squared threshold radius. The initial neuron activity is given by the following:

$$S_i(0) = \sum_{j=1}^{N} w_{ij} \geq 1, \tag{3}$$

which results in large initial activities for neurons inside a dense region. The "transmitting" process is captured with the following learning rule:

$$S_i(t+1) = S_i(t) + \alpha \sum_{j=1}^{N} w_{ij}(S_i(t) - S_j(t)) \tag{4}$$

where α is the learning rate parameter, often a time-decreasing function. If the activity of a neuron becomes negative during the transmitting process, the activity for that neuron is defined by $S_i(t) = 0$, removing it from the solution. As the transmitting process continues, the neurons at the periphery of the agglomeration surrender their activities to more centralized neurons. As the process runs its course, K neurons should remain distant from each other. The data points corresponding to these neurons are the cluster centers.

If the threshold is varied from a minimal value where each data point represents its own cluster center to a maximal value where all data points are contained in a single cluster, the algorithm explores the data set and provides insight into the number of clusters inherent to the data set.

One difference between the Litinskii-Romanov algorithm and the standard FOREL al-

gorithm is that the resultant cluster centers will always be drawn from the original data set, whereas in the standard FOREL algorithm the cluster centers are calculated based on the distribution of points within the threshold radius and can result in a cluster center that lies between the original data points.

Evaluation of Litinskii-Romanov Clustering

Due to the promising nature of the results, the Litinskii-Romanov clustering algorithm is evaluated for the point cloud segmentation problem. Its performance is first evaluated using common data sets and comparing the results of the developed implementation to those that Litinskii and Romanov report in [49].

The first data set used for comparison is comprised of 5 compact clusters, each with 50 data points. Each cluster of points is internally compact and the between cluster separation is large. The data for each cluster is generated using a Gaussian distribution. A plot of the data is provided in Fig. 11. A plot of the estimated number of classes versus threshold radius is shown in Fig. 12.

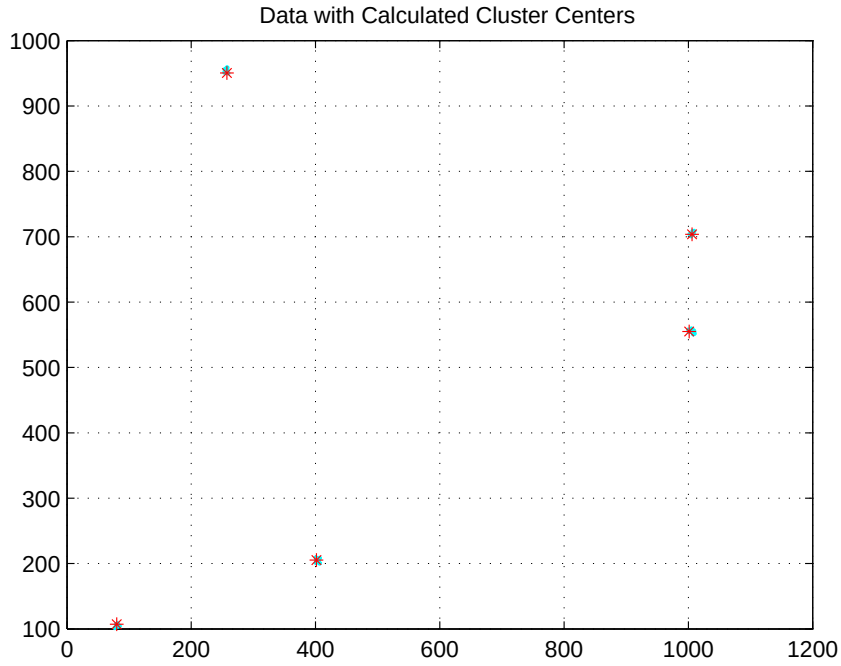


Figure 11. The input data of 5 compact clusters each containing 50 data points for the Litinskii-Romanov algorithm with cluster centers highlighted in red.

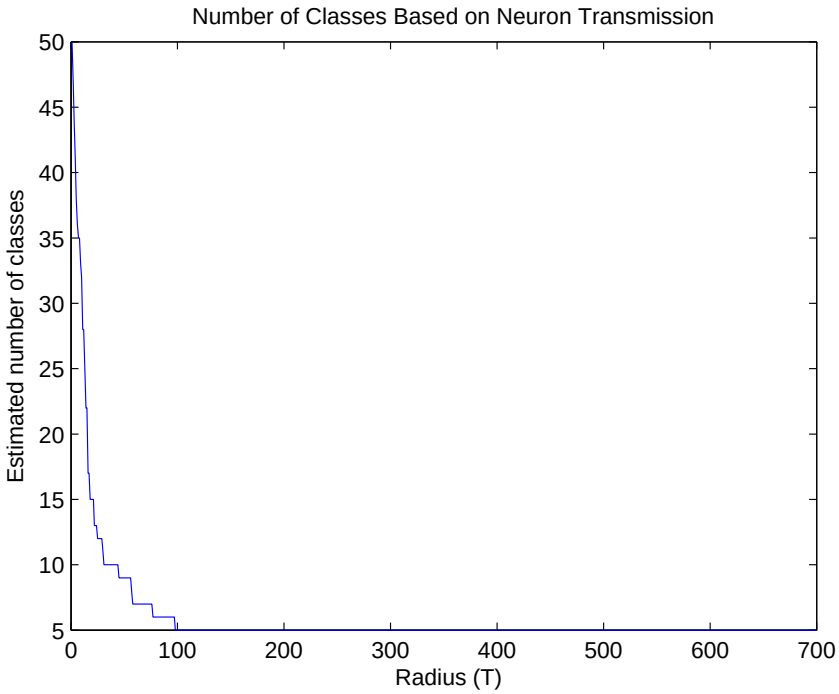


Figure 12. The Litinskii-Romanov algorithm's results of estimated number of clusters versus threshold radius plot for the 5 cluster data set.

The algorithm is also tested using the IRIS data set [25] to compare results with those published by Litinskii and Romanov in [49]. The resulting plot, shown in Fig. 13, shows some similarities with the published results; however, there are some significant differences. The published plot has a smoother and slower transition from 150 cluster centers to the expected 3 cluster centers (one for each of the three IRIS clusters), with the shape of the curve over this range of radii being concave down. The plateaus for 2 and 3 cluster centers are present in both plots, but the present results produce a plateau for 1 cluster at a much smaller radius than that in [49].

Applying the Litinskii-Romanov algorithm to the body scan data results in the cluster-vs-radius plot presented in Fig. 14. This plot shows the estimated number of clusters versus threshold radius. The terminal plateau of the plot indicates the data is comprised of two clusters. Though this is a viable result, it does not serve the intended purpose of correlating cluster centers to anthropometric marker locations. Since the goal is to locate the centers corresponding to the shoulders, hips, knees, ankles, elbows, wrists, and the head,

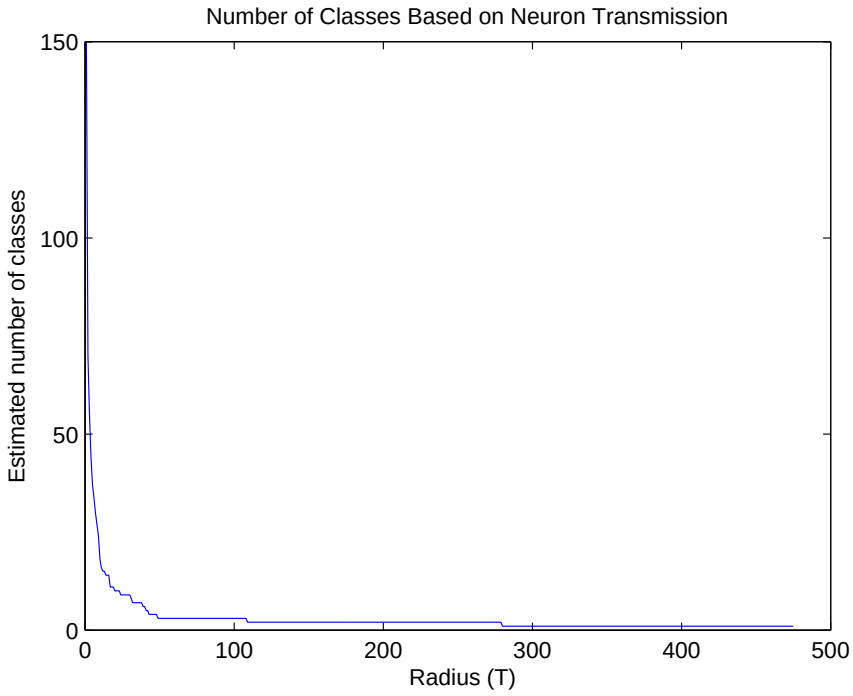


Figure 13. The Litinskii-Romanov algorithm's estimated number of clusters versus threshold radius plot for the IRIS data set [25].

an intermediate result is extracted corresponding to 13 cluster centers. The plot in Fig. 15 captures the cluster centers approximately 20% through the algorithm.

The Litinskii-Romanov algorithm performs well when working with easily separable clusters, as shown in the duplication of Litinskii and Romanov's results in [49]. The algorithm struggles in adequately providing anthropometrically useful cluster centers when applied to the body scan data. The general data exploration piece of the algorithm appears to find an appropriate number of clusters, but their class centers do not correspond to anthropometric marker locations.

2.3.3.3 Self-Organizing Map Clustering.

Though there is no literature directly addressing the application of applying a Self-Organizing Map (SOM) algorithm to 3-D body scan data, there is literature pertaining to the use of SOMs in reverse engineering mechanical components into CAD drawings. This is accomplished by taking LADAR scans of the object and then performing segmentation on

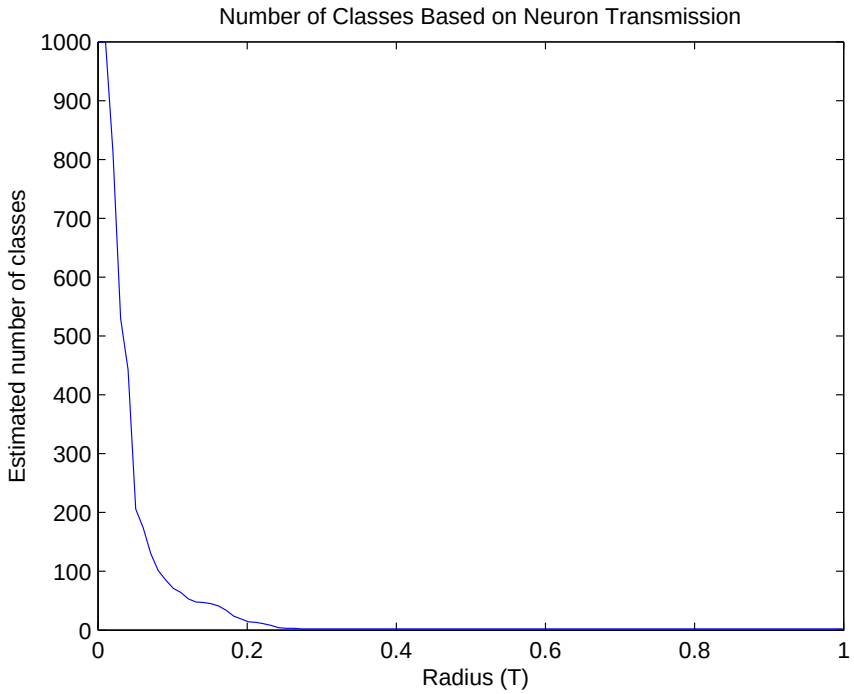


Figure 14. The Litinskii-Romanov algorithm's estimated number of clusters versus threshold radius plot for the body scan data.

the scans in order to extract the regions that comprise a surface of the object [51]. As part of the reverse engineering effort, there is a strong emphasis on performing edge detection on the point clouds in order to aid the segmentation process [87].

The SOM, an unsupervised clustering method, was introduced by Kohonen in 1982 [46]. In general, the goal of a SOM is the derivation of a feature map that captures the essential characteristics of the input data in a lower dimensional representation. The SOM network is comprised of two components: a connected lattice of neurons and a weight vector associated with each neuron. The lattice has an inherent connectivity and is a means to provide neighborhood preservation. A key feature of an SOM is that it preserves the inherent topology of the input data. The weight vectors are the learned component of the network [46].

The SOM algorithm begins with weight initialization. The three basic processes that comprise the formation of a SOM are as follows [35]: (a) competition, (b) cooperation, and (c) adaptation.

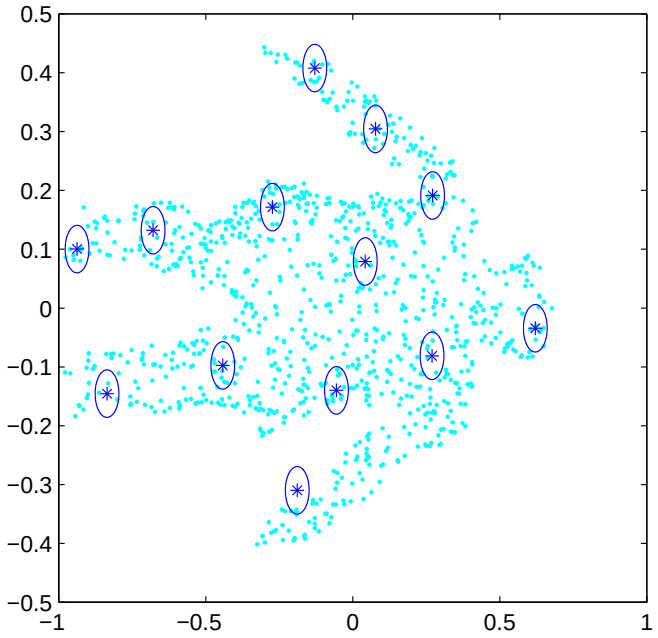


Figure 15. The calculated cluster centers at a stage 20% through the Litinskii-Romanov algorithm with taxons marked and representative threshold radius circles drawn around the taxon.

In the competition phase, each neuron calculates a value based on a discriminant function. The node with the smallest resulting value is declared the winning neuron. Based on the location of the winning neuron and a predetermined neighborhood function, a neighborhood of spatially related neurons become excited. The results of this excitement produce the adaption process in which the synaptic weights are adjusted to increase their value based on the input. Thus, similar inputs will provide enhanced responses to those particular lattice nodes [23].

In many applications, the discriminant function used to find the winning neuron compares the Euclidean distances such that the best-matching neuron is found as:

$$c = \arg\min_i \{\|x - m_i\|\} \tag{5}$$

where m_i is the weight of ith neuron. Updates to the winning neuron and neurons within

the neighborhood of the winner occur as:

$$m_i(t+1) = m_i(t) + h_{ci}(t)|x(t) - m_i(t)|, \tag{6}$$

where t is a discrete-time coordinate. The neighborhood function is given by $h_{ci}(t)$ which converges to zero as t approaches ∞ for Gaussian neighborhoods. Common neighborhood functions consist of a time diminishing Gaussian function and an n-connected components function in which 4- or 8- connected neighboring neurons are updated.

2.3.4 Geometry Computation.

The geometry computation section is a significant preprocessing step used in a segmentation from a surface approach. It is also a useful tool in evaluating the results of clustering algorithms. Presented here are discussions on Voronoi diagrams, Delaunay triangulations, and alpha hulls.

2.3.4.1 Voronoi Diagrams.

Voronoi tessellations are a common tool in computational geometry. In the simplest terms, a Voronoi region is defined in terms of its nearest neighbors in which the delineated region defines the region of points that are closer to that point than any other in the set of points [53].

More formally, given a set of points $\mathcal{X} = \{x_1, ..., x_n\}$ in $\mathbb{R}^d$, the Voronoi region of x_i is the set of points in $\mathbb{R}^d$ which are closer to x_i than to the other points in $\mathcal{X}$.

$$V_i = V(x_i; \mathcal{X}) = \{x \in \mathbb{R}^d : d(x, x_i) < \min_{j \neq i} d(x, x_j)\}. \tag{7}$$

The distance d is usually the Euclidean distance. The Voronoi tessellation, or Voronoi diagram, of the points is the set of points in $\mathbb{R}^d$ which have more than one nearest neighbor in $\mathcal{X}$ [53]. An example of a Voronoi diagram for a set of 20 random points is shown in Fig. 16.

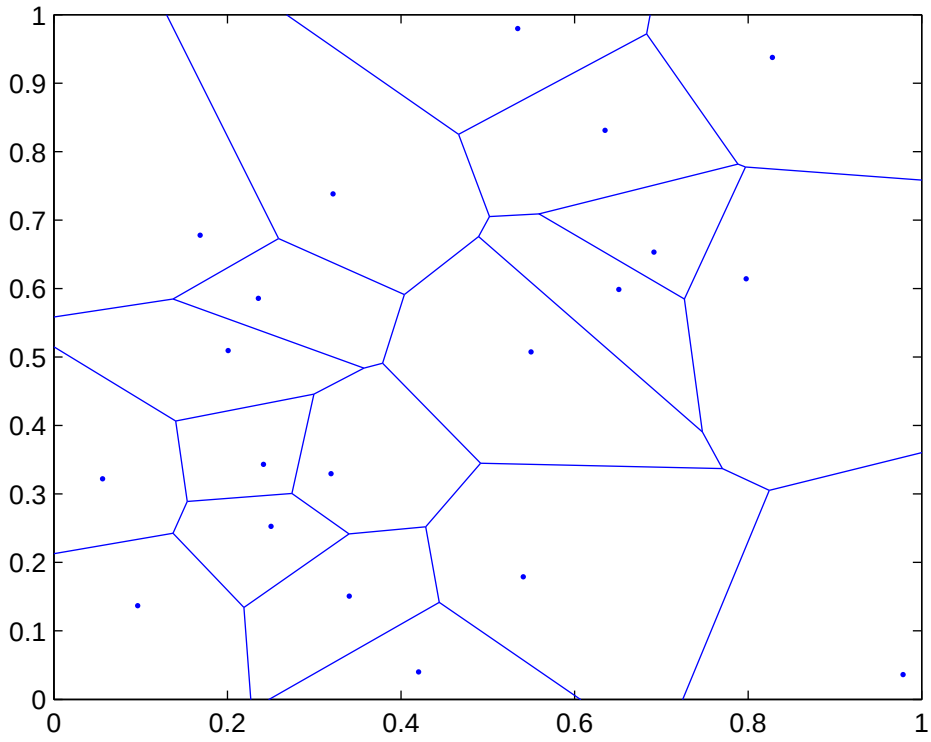


Figure 16. A Voronoi diagram for a set of 20 random points.

2.3.4.2 Delaunay Triangulation.

The Delaunay triangulation is the dual to the Voronoi tessellation. It is the graph whose vertices are $\mathcal{X}$, with an edge between x_i and x_j if and only if the Voronoi cells of x_i and x_j share an edge. The Delaunay triangulation theory dictates that for a set of points, P, there is a triangulation, $DT(P)$, such that no point in P lies within the circumscribed circle of any triangle in $DT(P)$ thereby minimizing the size of each of the triangles. A couple of other important properties of the Delaunay triangulation are that a convex hull is a subgraph of the Delaunay triangulation and no elements of $\mathcal{X}$ are contained within a triangle [53]. The computation time of a Delaunay triangulation is $\mathcal{O}(n \log n)$ [53]. An example of a Delaunay triangulation on the set of 20 random points used in the Voronoi tessellation example (Section 2.3.4.1) is shown in Fig. 17.

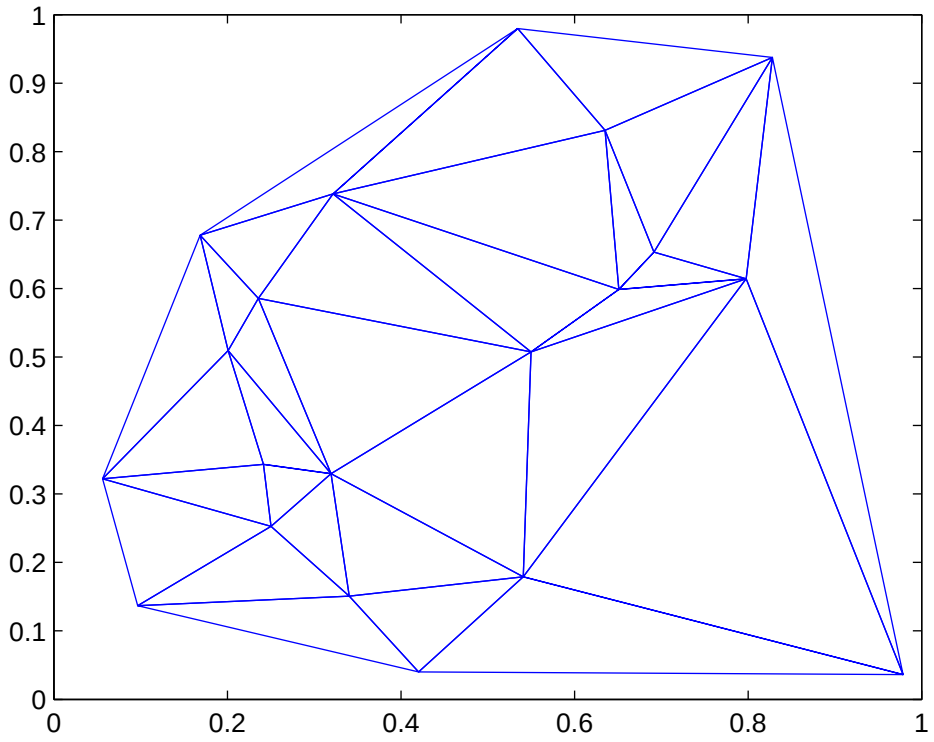


Figure 17. A Delaunay triangulation for a set of 20 random points (Fig. 16).

2.3.4.3 Alpha Hulls.

Alpha hulls are a generalization of the convex hull of a point set [20]. A convex hull for the random 20 point example is shown in Fig. 18. Marchette's definition of an alpha hull is presented as the following: for $\alpha > 0$, the alpha hull (also called the alpha shape) of a set of points $\mathcal{X} = \{x_1, ..., x_n\}$ is the graph with vertices $V = \mathcal{X}$ and edge set $E = \{v_i v_j : \exists D_\alpha$ with $v_i, v_j \in \partial D_\alpha$ and $\mathcal{X} \cap D_\alpha = \emptyset\}$, where D_α is a disk of radius α and ∂S denotes the boundary of the set S. For $\alpha < 0$, we require that the disk be of radius α, and $\mathcal{X} \cap D_\alpha = \mathcal{X}$. Finally, for $\alpha = 0$, the graph is the empty graph on $\mathcal{X}$ [53]. An example of an alpha hull for the set of 20 random points used in the Voronoi tessellation and Delaunay triangulation examples (Sections 2.3.4.1 and 2.3.4.2) is shown in Fig. 19.

A common description of the alpha hull process is to imagine rolling a disk of radius α around the set of points and the point pairs that can be touched by the disc of radius α

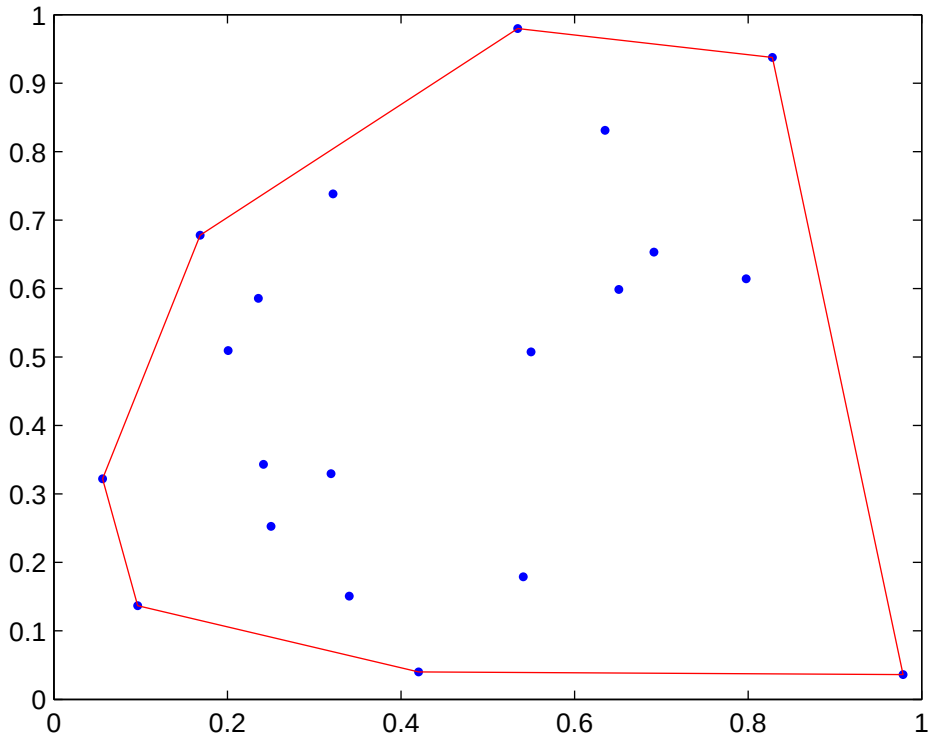


Figure 18. A convex hull for a set of 20 random points (Figs. 16, 17).

become a constituent line segment of the boundary that forms the shape of the data set. This concept is shown in Fig. 20. The effects of different values of α can be seen in Fig. 21 in which the appearance of the resulting alpha hull varies greatly based on the selected α. The result of $\alpha = \infty$ is the convex hull.

2.3.5 Skeletonization.

In segmentation from a surface approaches, skeletonization acts as the segmentation method whereas it serves as a post-processing step for segmentation from point cloud approaches. Discussed here are the concepts of Reeb graphs and medial axes.

2.3.5.1 Skeletonization from a Surface.

Background on Skeletonization

Due to the significant number of papers that addressed segmentation from a polygo-

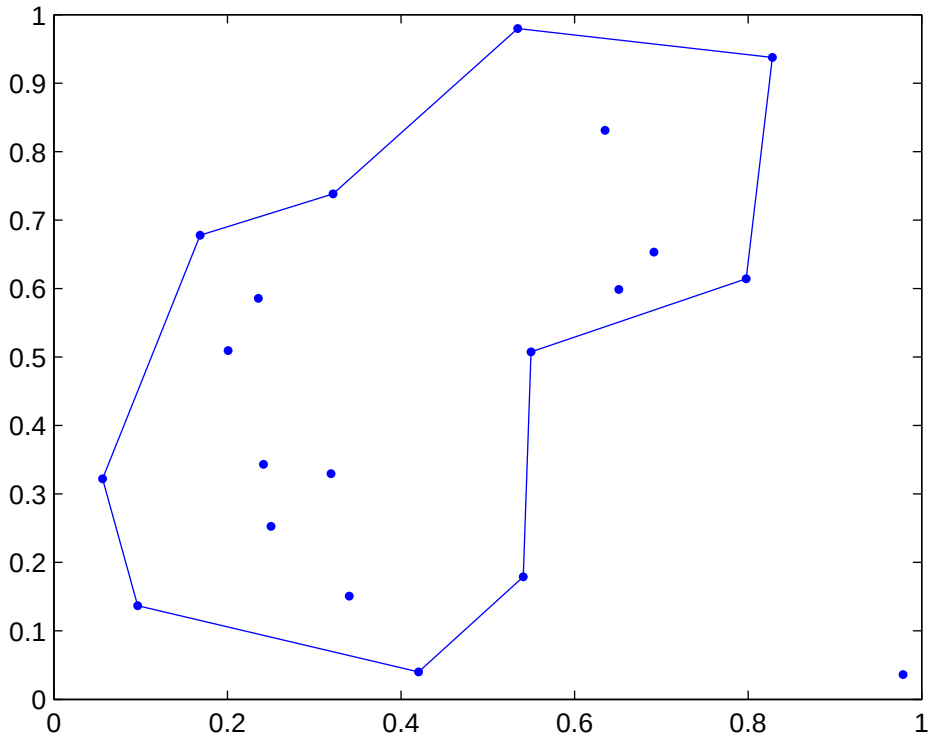


Figure 19. An alpha hull with $\alpha = 0.25$ for a set of 20 random points (Figs. 16, 17, 18).

nally meshed surface, the approach by [92] is described. There are a variety of methods available to generate a polygonal mesh from a point cloud, but Delaunay triangulation is a common method used in computational geometry, mathematics, and computer graphics and is selected for evaluation. It is a similar algorithm to the Voronoi tessellation algorithm used for creating equidistant decision boundaries for classification problems [1]. A Delaunay triangulation mesh extracted from the 3-D point cloud data is shown in Fig. 22.

Once a surface is generated, the skeleton is derived. A method drawn from 2-D image processing is the Medial Axis Transform [41]. It was formalized by Harry Blum of the Air Force Research Laboratories in Cambridge, Massachusetts, in which he defined a medial axis for computing a 2-D skeleton using a "grass field fire propagation" model to describe the resultant line segment generated if a fire was set at the boundary of a shape and was allowed to run to the center of the shape [7]. This description serves as a basic understanding for the analogous mathematical definitions further developed under Reeb graph and contour

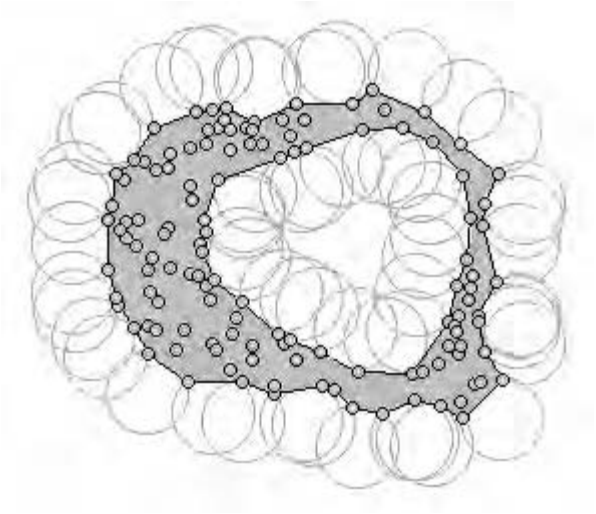


Figure 20. A standard representation of the generation of an alpha hull by rolling a disk of radius α around the set of points.

tree theory in computational geometry and topology [4].

The Reeb Graph, a mathematical construct based on Morse theory, represents the topological skeleton of a multi-dimensional object. A height function is used to calculate level-set curves which represent a vertex in a graph, producing a skeleton of the object [19]. Implementations of a basic Reeb Graph are generally sensitive to noise in the data. As such, the numerous adaptations of this approach include the following: examining various height functions [81], using a discrete Reeb Graph [86, 92], and affine invariant Reeb graphs [62].

A simplified description of Reeb graph theory is that for a surface, $DT(P)$, such as produced from a Delaunay triangulation, a height function, f, can be defined such that the contour of the surface is described by a skeleton, the reduction of the shape to its minimal form. The skeleton is derived in such a way that it connects all of the critical points of the object. An example of the output of a Reeb graph algorithm is shown in Fig. 23 (from [89]).

An example of a skeleton derived from a 3-D point cloud is shown in Fig. 24 for the lower body portion (waist and below) of the point cloud data. In this example, a medial axis calculation is used. The medial axis is found by minimizing the Euclidean distance of each point on the surface with those contained within the surface.

Evaluation of a Skeletonization from a Surface Algorithm

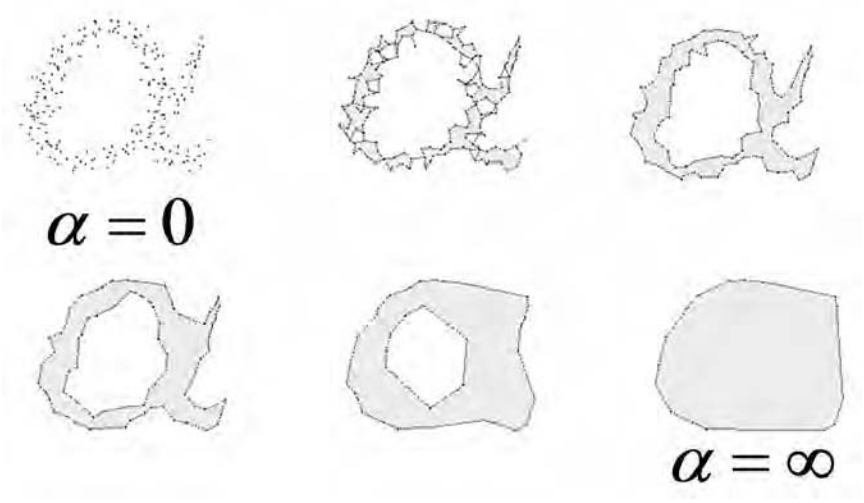


Figure 21. A standard depiction of the effect of varying the size of α on the resulting alpha hull from a set of points.

Computational adversities, similar to those faces in the LLE implementation, are encountered with the skeletonization process. In part, this increased processing time is due to the additional step of generating a surface from the point cloud, though the skeletonization algorithms are computationally intensive unto themselves.

The published results for skeletonization from a meshed surface show promising results for extracting a skeleton from the point cloud data [9, 31, 89, 92]. However, attempts to duplicate those efforts did not produce satisfactory results. Additionally, once a skeleton is obtained, marker estimation would need to be extrapolated from the critical points of the surface as per the Reeb graph theory approach, or some other method of marker extraction would need to be developed.

2.3.6 Extracting Marker Locations.

2.3.6.1 Incorporating Anthropometric Information.

A topic that is not addressed in the available literature on segmenting human body scans is how to extract anthropometric markers from the resulting segmentation. An interesting approach to the problem is based on historical research into the proper proportioning of the human body. A famous example of this research is Leonardo da Vinci's classic Vitruvian

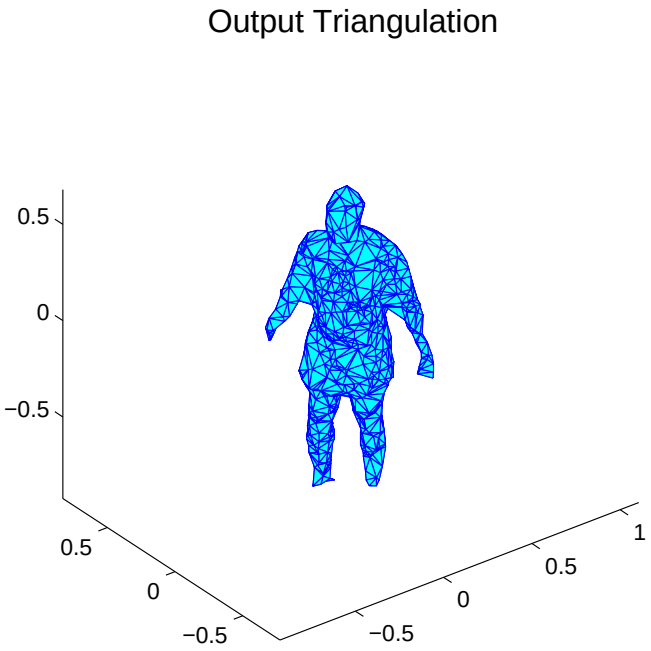


Figure 22. A triangulated surface mesh applied to a 3-D point cloud. The mesh is derived using a Delaunay triangulation based algorithm.

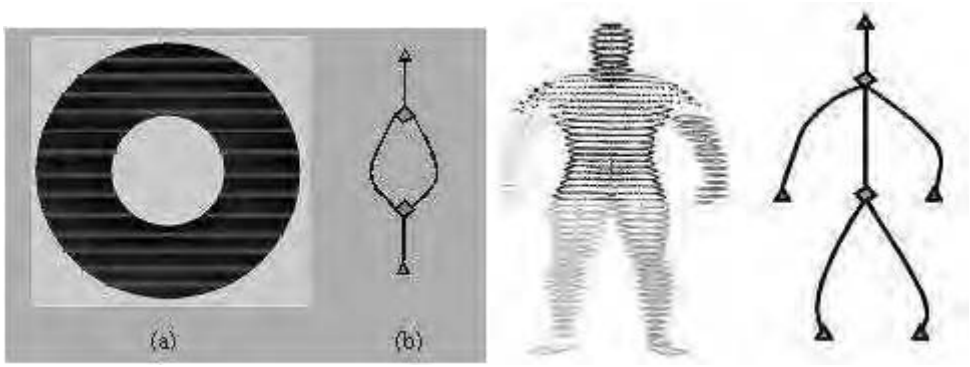


Figure 23. A Reeb graph of a torus and a Reeb graph of a human body shape as produced in [89].

Man drawing from approximately 1487 [17]. The accompanying text, known as the **Canon of Proportions** delineated a series of proportions for body parts. This work was brought to fruition in the 19th century by the German scientist, Adolf Zeising [63]. His work was based on the principle of the golden section. Given a line segment, shown in Fig. 25, defined by endpoints A and B, there exists some point C on this line segment that divides the line

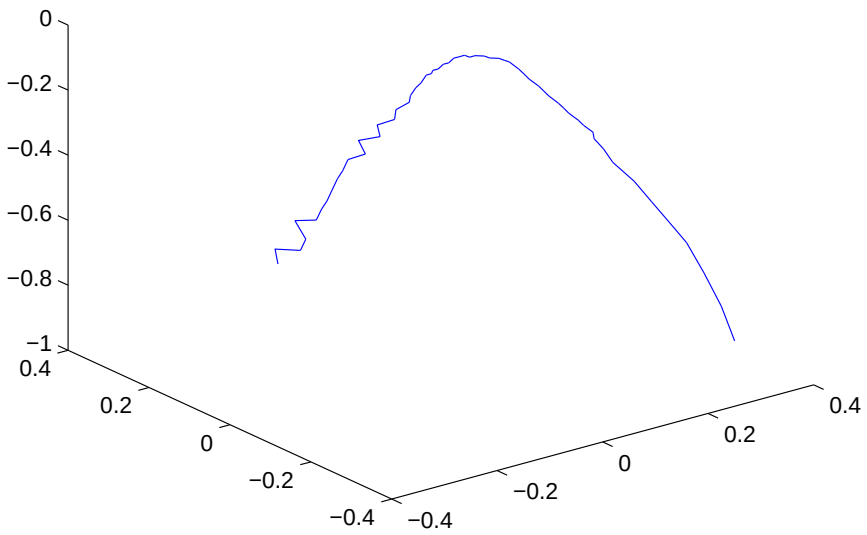


Figure 24. A skeleton of the lower body derived from the triangulated surface mesh. It is based on a medial axis algorithm to find the central set of points equidistant from the surface.

such that:

$$\frac{AB}{AC} = \frac{AC}{CB} \tag{8}$$

$$(AC)^2 = (AB)(CB). \tag{9}$$

If the length of line segment AC is defined as M and the length of line segment CB is defined as m and substituted into Eqns. 9, then

$$\frac{M+m}{M} = \frac{M}{m} \tag{10}$$

$$M^2 = m(M+m) \tag{11}$$

Solving the resulting quadratic equation for the ratio, M/m, provides a single positive root of 1.6180. From this result, and the proper selection of the points, the location of body markers are obtained. The initial step is to let AB represent the total height of the individual and point C represent the location of the umbilicus. This golden section ratio

can then be used to further partition the body into subsequently smaller sections [63].

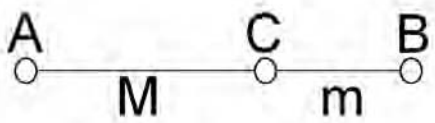


Figure 25. A line segment depicting the relationship the points A, B, and C and the lengths M and m for the purposes of illustrating the concept of the golden section as described in [63].

Unfortunately, in practice this systematic proportioning of the human body fails. The length of the upper arms is greater than that predicted by the model, the arms are longer than expected, the biacromonial diameter is greater than predicted, the length of the bust is lower than calculated, and the lower limbs are completely different from expected. Furthermore, the height of the leg is longer than expected, the shank is longer and the knee is located higher than expected. In the end, the golden section ratio may provide a rough estimate of marker locations, but the true locations of the markers are not accurately predicted by the golden section ratio method. Or rather, the "ideal" proportions do not provide a good prediction of the wide variance of human proportions [63]. However, incorporating such information may prove useful in locating markers in the LADAR data.

2.4 Computing Truth Markers

2.4.1 Exploitation of the CAESAR LADAR Data.

Work accomplished by the Air Force Research Laboratory's 711th Human Performance Wing in gender classification is based on the goal of "identifying anthropometric measurement combinations that best distinguish between genders in order to address issues related to apparel and workstation design and performance [27]." The results from this research serve as the baseline performance measures used in evaluating the segmentation algorithm developed in this work. The research in [27] combined traditional 1-D measurements with additional information obtained from the 3-D body scans. This combination provides 97 anthropometric measures from which to discriminate gender. In the final analysis, only 16

Table 1. Significant measures for the gender classification problem in [27]. Values shown are the means (standard deviations) for the anthropometric measurements given in mm or kg.

	North America		Italy		The Netherlands	
anthropometric measure	males	females	males	females	males	females
a) waist back	481 (35)	399 (28)	468 (30)	405 (30)	475 (36)	401 (29)
b) bustpt brth	236 (27)	187 (23)	217 (20)	183 (21)	230 (25)	206 (26)
c) hip brth, sitting	382 (36)	408 (46)	359 (25)	375 (29)	382 (29)	416 (38)
d) ankle circ	269 (15)	240 (15)	263 (13)	238 (12)	267 (16)	246 (16)
e) bi-lat fem epi brth sit	463 (55)	361 (64)	449 (46)	350 (48)	427 (59)	340 (53)
f) chest circ	1040 (109)	958 (124)	958 (79)	890 (80)	1015 (102)	998 (119)
g) chest grth at scye	1055 (96)	921 (100)	978 (71)	854 (60)	1022 (85)	943 (88)
h) neck base circ	468 (30)	410 (27)	475 (21)	425 (21)	489 (36)	441 (33)
i) weight	86 (18)	69 (18)	73 (11)	58 (9)	84 (16)	73 (16)
j) triceps skinfold	13 (7)	24 (10)	13 (7)	21 (7)	10 (5)	19 (8)
k) bi-lat hum epi brth sit	561 (53)	475 (57)	557 (45)	469 (39)	562 (49)	496 (49)
l) shoulder brth	496 (36)	430 (35)	459 (27)	405 (23)	472 (29)	431 (31)
m) radiale-stylion lth rt	265 (17)	237 (15)	265 (14)	239 (14)	265 (17)	239 (16)
n) waist circ (pref)	914 (125)	789 (135)	843 (83)	752 (78)	918 (109)	845 (131)
o) knee height	562 (31)	509 (28)	541 (26)	497 (24)	558 (35)	514 (28)
p) sitting height	926 (40)	865 (36)	908 (35)	855 (30)	945 (41)	887 (38)

measures were significant to the discriminant functions as listed in Table 1.

Based on the use of the above features, the authors in [27] achieved classification accuracy of 99.5% for the North American and Italian databases and 98.5% for the Netherlands database. Further evaluation of the discriminators indicated that two measures were of greater significance. These are the sitting hip breadth and the sitting bi-lateral femoral epicondyle breadth, or more commonly, the sitting knee breadth. These measures cause a problem when attempting to apply a classification algorithm based on those features to a data set consisting entirely of standing subjects. It is impossible to accurately calculate these features from a scan of a standing subject. The benefit of being able to use these two features in discriminating gender is that there are distinct differences between males and females in these two areas. The female's pelvis is larger and broader than that of a male's. A male's pelvis is taller, narrower, and more compact than a female's [78]. Women tend to be proportionately larger in the hips and smaller in knee breadth than men [27]. A contributing factor to this observation is that women tend to store more adipose tissue in the gluteal-femoral region, whereas men tend to store fat around the abdomen [6].

Table 2. Truth marker set.

Set of Calculated Truth Markers
Left ankle center
Right ankle center
Left knee center
Right knee center
Left hip center
Right hip center
Left wrist center
Left elbow center
Left shoulder center
Right wrist center
Right elbow center
Right shoulder center
Head center

Table 3. Feature set used for classification.

Feature Set
Left shank length
Right shank length
Left thigh length
Right thigh length
Left lower arm length
Right lower arm length
Left upper arm length
Right upper arm length
Shoulder breadth
Hip breadth
Hip height
Height

The marker set calculated by the segmentation algorithm consists of the markers presented in Table 2. Since the problem addressed in this thesis contains only the 3-D LADAR data, the feature set was restricted to only those measurements obtainable from the calculated anthropometric markers. Therefore, measurements of weight, triceps skinfold, and measurements obtained from sitting subjects are removed from the feature set. This results in a reduced feature set used for the classification process. The reduced feature set is presented in Table 3.

2.4.2 Truth Marker Processing.

The CAESAR data set consists of a 3-D LADAR body scan file as well as a landmark file that consists of the locations of 72 markers affixed to the subject's body. The markers that the tester places on the subject is shown in Fig. 26. As is the nature of obtaining true marker locations, there are some errors associated with placing the markers on the skin since the flesh is not tightly attached to the bone. Additionally, the tester's skill in placing the markers is important since many of the markers are attached to either side of the desired marker location. For example, in determining the knee center location markers are affixed at the femoral epicondyle medial and femoral epicondyle lateral locations as shown in Fig. 27 [72]. The desired marker thus becomes the average of those two markers. Of course, if one of the markers is placed too high or too low, the resulting calculated knee center will not correspond with the true knee center location.

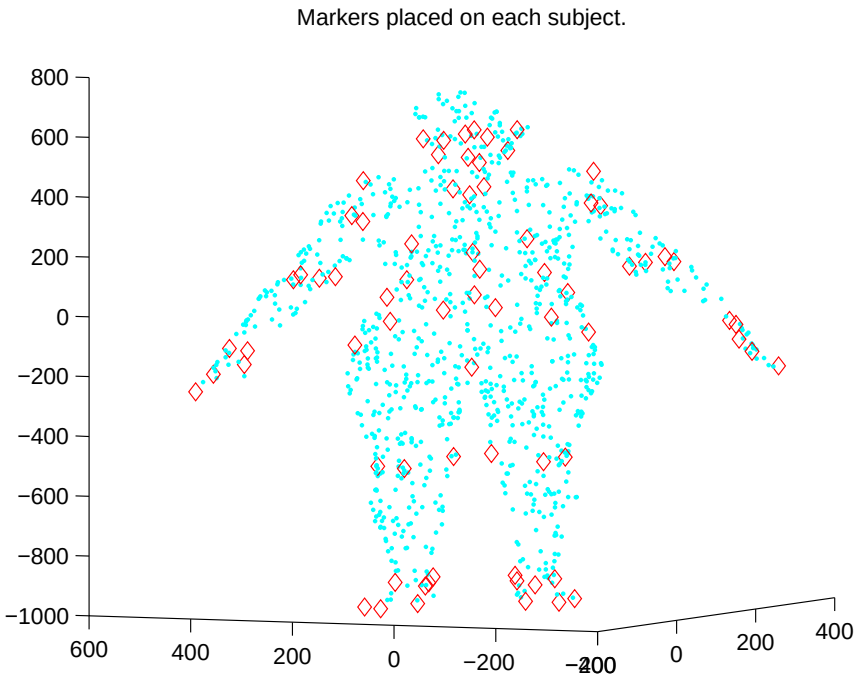


Figure 26. Plot of the markers placed on the subject. For clarity, the markers are depicted as red diamonds.

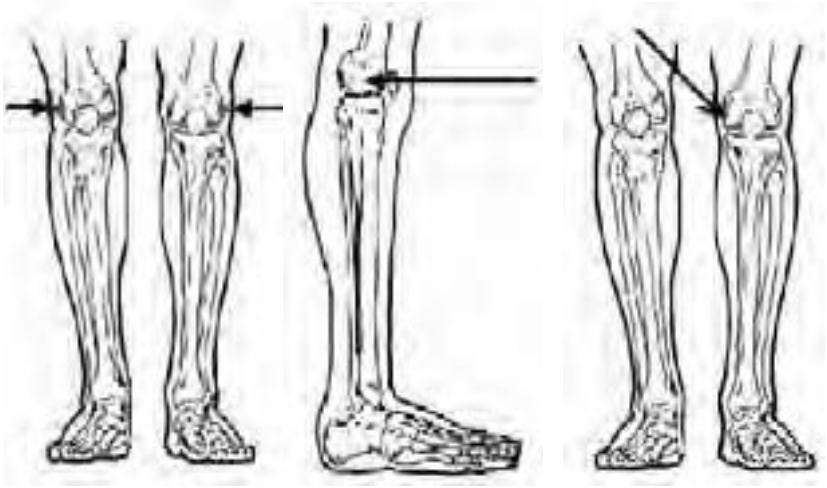


Figure 27. Line drawing depiction of the femoral epicondyle lateral (left and center) and medial (right) marker placement. The drawings are from [72].

2.4.2.1 Hip Joint Center (HJC) Calculation.

Finding the location of the hip joint center (HJC) is a difficult problem since by its location it is impossible to place a marker on the skin to completely describe the location of the HJC. In general, the only sure method of accurately locating the HJC is through x-ray or MRI scans [42]. Significant research in calculating the HJC exists in the medical, anthropometry, and kinesiology communities. Several methods with a high degree of HCJ accuracy rely on placing the markers at the anterior superior iliac spines (ASISs) and the posterior superior iliac spines (PSISs), as shown in Fig. 28 [10].

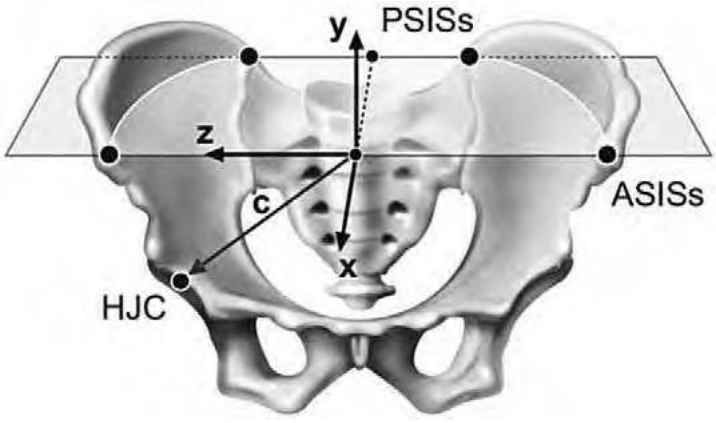


Figure 28. Pelvic anatomical frame showing ASISs, PSISs, and the local reference system [10].

As seen in Fig. 28, the z-axis is defined as the line passing through the ASISs. The HJC lies on the vector $\vec{c}$. To produce the HJC, the process begins by creating the z-axis for a coordinate system local to the hips by subtracting the midpoint between the left and right ASIS markers from the right ASIS marker location:

$$\vec{z} = p_{ASIS_R} - p_{ASIS_C}. \tag{12}$$

A temporary vector that lies in the plane depicted in Fig. 28 is created by subtracting the midpoint between the left and right ASIS markers from the midpoint between the left and right PSIS markers:

$$\vec{t_a} = p_{PSIS_C} - p_{ASIS_C}. \tag{13}$$

The y-axis of the coordinate system is found by taking the cross product of the two vectors:

$$\vec{y} = \vec{t_a} \times \vec{z}. \tag{14}$$

The x-axis of the right-handed coordinate system is found by taking the cross product of the first two axes:

$$\vec{x} = \vec{y} \times \vec{z} \tag{15}$$

In the defined local coordinate system, the HJC can be found as [10].

$$RHJC_{Local} = [0.36PW, -0.22PW, 0.3PW], \tag{16}$$

$$LHJC_{Local} = [-0.36PW, -0.22PW, 0.3PW], \tag{17}$$

where PW is the pelvic width defined as the Euclidean distance between the ASIS markers. Once the HJC is computed, it is transformed into the global coordinate system.

The axis vectors constitute a rotation matrix used to associate the location of the HJC in the local coordinate system to its location in the global coordinate system. The rotation

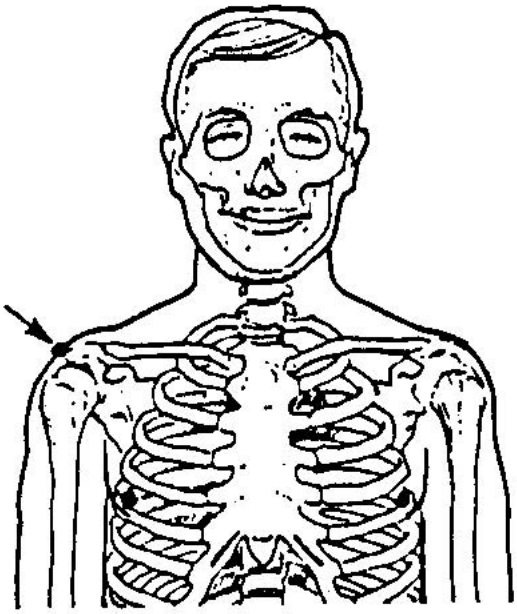


Figure 29. Line drawing depicting the location of the placement of the acromion marker prior to the body scan. Note its proximity to the shoulder joint center. Drawing is taken from [72].

matrix is described by:

$$\mathbf{R}_{HJC} = \begin{bmatrix} z_1 & y_1 & x_1 \\ z_2 & y_2 & x_2 \\ z_3 & y_3 & x_3 \end{bmatrix} \tag{18}$$

The HJC in the global coordinate system is then described as:

$$HJC_{Global} = p_{ASIS_C} + HJC_{Local}\mathbf{R}_{HJC}^{T}. \tag{19}$$

2.4.2.2 Shoulder Joint Center (SJC) Calculation.

The shoulder joint center (SJC) calculation is addressed in a similar manner to that described for the HJC [90]. However, in practice good results can be achieved by calculating the SJC based on a standard offset from the acromion marker location [26, 70]. The correspondence of the acromion marker to the SJC can be seen in Fig. 29 [72].

In order to examine the results of the joint center calculations, the combination of all of the markers is presented in Fig. 30. Only 0.5% of the 3-D body scan data points are shown to provide sufficient clarity to see the marker locations while retaining the scanned subject's body structure. The landmark file is used in conjunction with the marker determination algorithm described in Sec. 2.4 to calculate the standard set of anthropometric markers

used in this research as listed in Table 2.

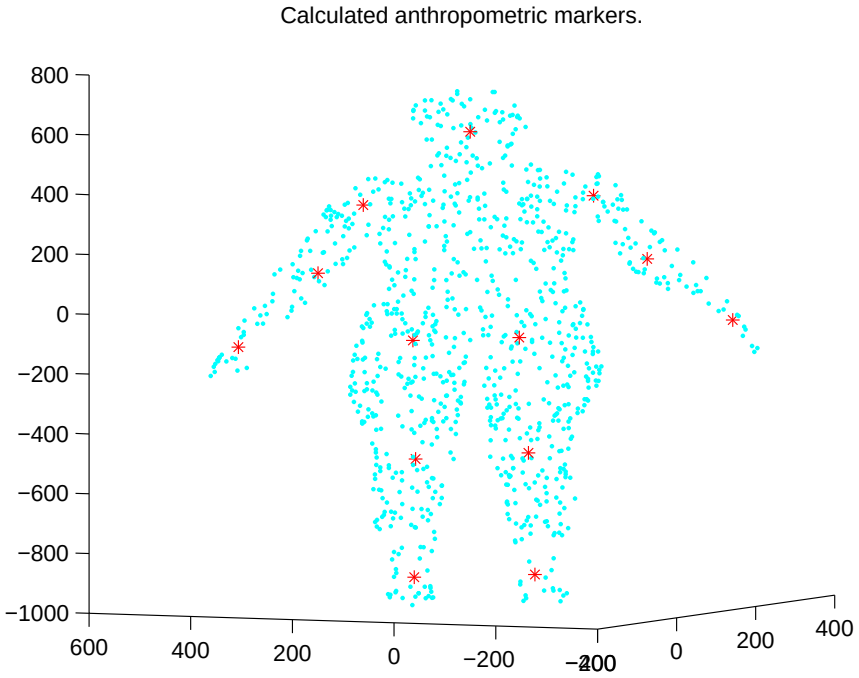


Figure 30. Plot of the anthropometric truth markers calculated from the markers placed on the subject during the scan. Only 0.5% of the 3-D point cloud data is plotted to provide clarity.

2.5 Evaluating the Marker Estimation Algorithm

In order to evaluate how well the marker estimation algorithm performs, we need a method to quantify performance. One way of doing this is to use the classification results described in [27].

2.5.1 Classification.

The classification problem can be accomplished through a variety of methods, and through the course of this research effort several methods were explored and tested. However, since the sponsor obtained the performance benchmark using linear discriminant analysis, it is the method applied to the results presented in this thesis.

2.5.2 Linear Discriminant Analysis.

Linear discriminant analysis (LDA) is a method to find a linear combination of features with which to characterize two or more classes. Define the prior probability of class k as π_k with $\sum_{k=1}^{K} \pi_k = 1$. The class-conditional probability density function of X in class $G = k$ is $f_k(x)$. From decision theory, optimal classification is based on the class posteriors $Pr(G|X)$. Calculating the posterior probability based on an application of Bayes theorem results in

$$Pr(G = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{l=1}^{K} f_l(x)\pi_l}. \tag{20}$$

By maximum *a posteriori* decision theory, the optimal classification is based on the class posteriors:

$$\begin{aligned} \hat{G}(x) &= \arg\max_k Pr(G = k|X = x) \\ &= \arg\max_k f_k(x)\pi_k. \end{aligned} \tag{21}$$

If each class density is modeled as a multivariate Gaussian having the form:

$$f_k(x) = \frac{1}{(2\pi)^{p/2}|\mathbf{\Sigma}_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)^T \mathbf{\Sigma}_k^{-1}(x-\mu_k)}. \tag{22}$$

then

$$\begin{aligned} \hat{G}(x) &= \arg\max_k Pr(G = k|X = x) \\ &= \arg\max_k f_k(x)\pi_k. \end{aligned} \tag{23}$$

When it is assumed that the classes have a common covariance matrix, $\mathbf{\Sigma}_k = \mathbf{\Sigma}, \forall k$, the

special case of linear discriminant analysis (LDA) is obtained.

$$\begin{aligned}
\hat{G}(x) &= \arg\max_k \log(f_k(x)\pi_k) \\
&= \arg\max_k [-\log((2\pi)^{p/2}|\mathbf{\Sigma}|^{1/2}) - \frac{1}{2}(x-\mu k)^T\mathbf{\Sigma}^{-1}(x-\mu k) + \log(\pi_k)] \\
&= \arg\max_k [-\frac{1}{2}(x-\mu k)^T\mathbf{\Sigma}^{-1}(x-\mu k) + \log(\pi_k)] \\
&= \arg\max_k [x^T\mathbf{\Sigma}^{-1}\mu_k - \frac{1}{2}\mu_k^T\mathbf{\Sigma}^{-1}\mu_k - \frac{1}{2}x^T\mathbf{\Sigma}^{-1}x + \log(\pi_k)] \\
&= \arg\max_k [x^T\mathbf{\Sigma}^{-1}\mu_k - \frac{1}{2}\mu_k^T\mathbf{\Sigma}^{-1}\mu_k + \log(\pi_k)]. \qquad (24)
\end{aligned}$$

The linear discriminant function is defined as:

$$\delta_k(x) = x^T\mathbf{\Sigma}^{-1}\mu_k - \frac{1}{2}\mu_k^T\mathbf{\Sigma}^{-1}\mu_k + \log(\pi_k), \qquad (25)$$

resulting in the following classification rule:

$$\hat{G}(x) = \arg\max_k \delta_k(x). \qquad (26)$$

Looking at the two class case where the classes are represented by k and l, the decision boundary is described by

$$\{x : \delta_k(x) = \delta_l(x)\} \qquad (27)$$

and the following holds:

$$\log\frac{\pi_k}{\pi_l} - \frac{1}{2}(\mu_k+\mu_l)^T\mathbf{\Sigma}^{-1}(\mu_k-\mu_l) + x^T\mathbf{\Sigma}^{-1}(\mu_k-\mu_l) = 0. \qquad (28)$$

This results in a decision boundary linear in x. In practice, the parameters for the Gaussian distribution are unknown, so the parameters can be estimated from the training data as

follows:

$$\hat{\pi}_k = \frac{N_k}{N}, \tag{29}$$

$$\hat{\mu}_k = \sum_{g_i=k} \frac{x_i}{N_k}, \text{and} \tag{30}$$

$$\hat{\mathbf{\Sigma}} = \sum_{k=1}^{K} \sum_{g_i=k} (x_i - \hat{\mu}_k)(x_i - \hat{\mu}_k)^T/(N-K), \tag{31}$$

where N_k is the number of class-k observations [34].

If the covariance matrices are not assumed to be equal, then the quadratic elements in x remain. This establishes the quadratic discriminant function (QDA) [5] as:

$$\delta_k(x) = -\frac{1}{2}\log|\mathbf{\Sigma}_k| - \frac{1}{2}(x-\mu_k)^T\mathbf{\Sigma}_k^{-1}(x-\mu_k) + \log\pi_k. \tag{32}$$

The classification rule is

$$\hat{G}(x) = \arg\max_k \delta_k(x) \tag{33}$$

and the decision boundaries are quadratic in x.

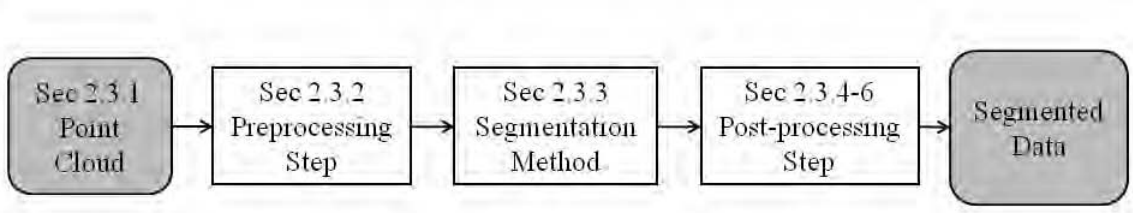
2.6 Summary

This chapter reviewed related literature dealing with segmenting point cloud data, introduced a conceptual paradigm in which to approach the marker estimation problem, and described various tools that will be used in the next chapter to describe the segmentation algorithm's flow. The next chapter goes into depth on the particular process used to estimate marker locations. Based on promising concepts in approaches from the literature review and preliminary algorithm evaluation, research into the development of the segmentation algorithm focuses on building it around the Self-Organizing Map. A description of the theory behind the individual components is provided in this section beginning with a few computational geometry concepts. The integrating of these methods to produce anthropometric markers is covered in Chapter 3.

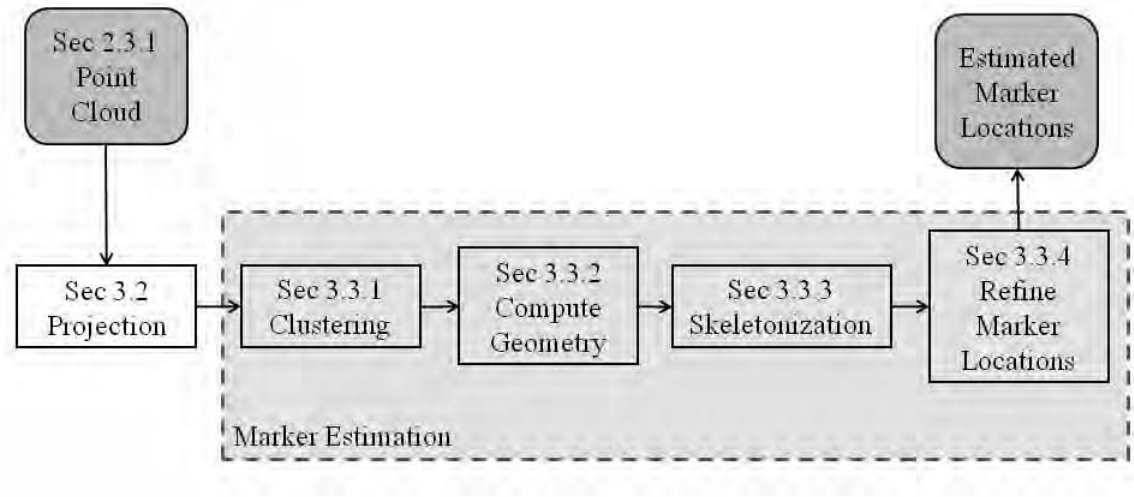
3. Methodology

This chapter provides the methodology that serves as the initial foray into extracting marker locations from 3-D LADAR point cloud data. The segmentation directly from the point cloud (Section 2.2.1) is chosen for the general approach and a description of the overall process is depicted in Fig. 31.

The decision to proceed with the segmentation from point cloud approach vice the segmentation from a surface approach is based on several factors. The primary reason is that, since the end goal is to derive anthropometric markers, the surfacing approach would



(a) Fundamental diagram



(b) Segmentation from a point cloud diagram

Figure 31. Diagrams of (a) the fundamental flow introduced in Section 2.3 and (b) the overall segmentation process specific to the segmentation from a point cloud approach. The 3-D point cloud data is the input to the system, and the estimated marker locations are the product of the marker estimation algorithm.

require an additional step to appropriately locate the markers. Furthermore, segmenting from a point cloud is more computationally efficient. The surfacing and medial axis of a 3-D point cloud are more computationally intensive than projection and clustering often used in segmenting point clouds.

In relating Fig. 31b to the conceptual paradigm shown in Fig. 31a, the preprocessing block in the current diagram consists of two steps: (a) scaling and (b) projection. The segmentation block of Fig. 31a is clustering in Fig. 31b. The post-processing step of Fig. 31a consists of (a) computation of the geometry, (b) skeletonization, and (c) marker refinement.

3.1 Algorithm Overview

The general concept of the segmentation from point cloud approach is to project the 3-D point cloud data into a meaningful dimensionally reduced space on which a clustering algorithm extracts cluster centers from the distribution of the projected data. The calculated cluster centers are then refined to better estimate anthropometric marker locations. Based on the background research discussed in Sections 2.3.4, 2.3.5, and 2.3.6, the cluster center refinements depend on incorporating anthropometric information into the algorithm. The rest of the chapter details the choices made for each of the blocks in Fig. 31b.

3.2 Projection

Due to the nature of the feature set used in the classification problem, namely a feature set of body segment lengths as described in Sec. 2.4.1, the dimensionality of the data is reduced to simplify the complexity of the problem. A good choice for dimensionality reduction is one that provides a coronal plane view of the subject as shown in Fig. 7. The coronal plane provides a complete view of the markers we are trying to calculate since, for a standing subject, the anthropometric marker set we desire is roughly situated in a single plane. As discussed in Section 2.3.2.1, PCA provides such a result. An example of the resulting projection is shown in Fig. 32. The data is projected into two dimensions using the first two principal components.

In the data preprocessing phase, the data are also scaled. The primary reason for scaling the data is to make use of the results found in Section 2.3.6 about locating a region in which we can expect a marker to lie based on the golden section ratio. Working in a scaled space minimizes the effects of the wide variations in human physiques, standardizing the region in which markers are expected to exist. The calculated markers are rescaled at the end of the marker extraction algorithm.

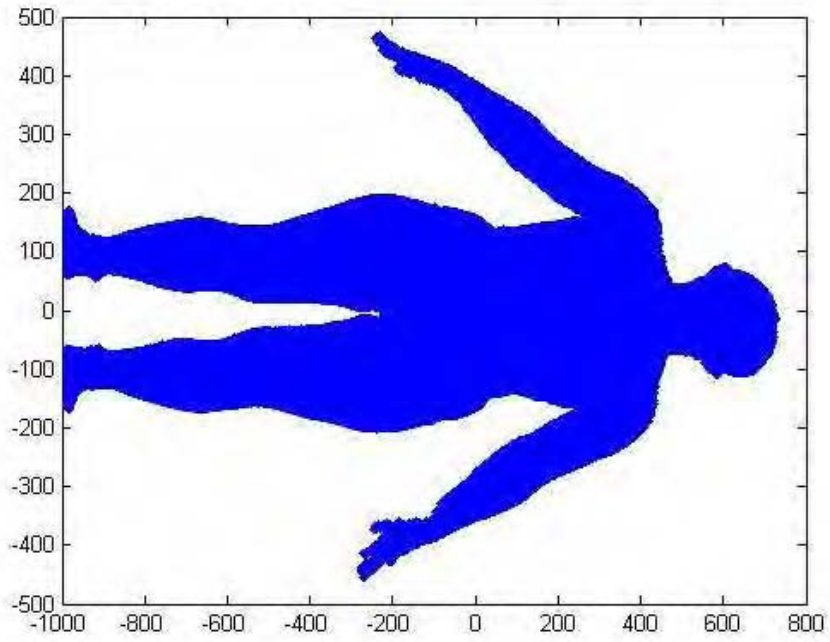


Figure 32. Point cloud data from [72] projected into 2-D using the first two principal components.

3.3 Marker Estimation Algorithm

The marker estimation phase of the methodology consists of four parts: (a) clustering, (b) computing a bounding geometry, (c) skeletonization, and (d) marker refinement. This section describes the methods employed and the reasons for their selection.

3.3.1 Self-Organizing Map.

The clustering method chosen is the Self-Organizing Map (SOM). The first reason for this choice is its performance in mapping the distribution of the projected point cloud data. Secondly, the connectivity of the lattice structure is capable of imparting more information than the K-Means or FOREL approaches. A detrimental aspect of both the SOM and K-Means approaches is the possibility of neurons or means falling outside the projected data set. The FOREL method on the other hand constrains its taxons to an actual data point, hence avoiding the problem. A method for dealing with the dead neurons, as the outlying neurons are called in SOMs, is presented in the next section.

The algorithm starts with the 2-D projected point set. From this set of points, the SOM algorithm is run. The algorithm uses a 22-node 1-D lattice. This lattice size was determined empirically through the analysis of the quantization and topographic errors. A representation of the typical output of the SOM portion of the segmentation algorithm is shown in Fig. 33. The rationale behind these selections is provided in Chapter 4.

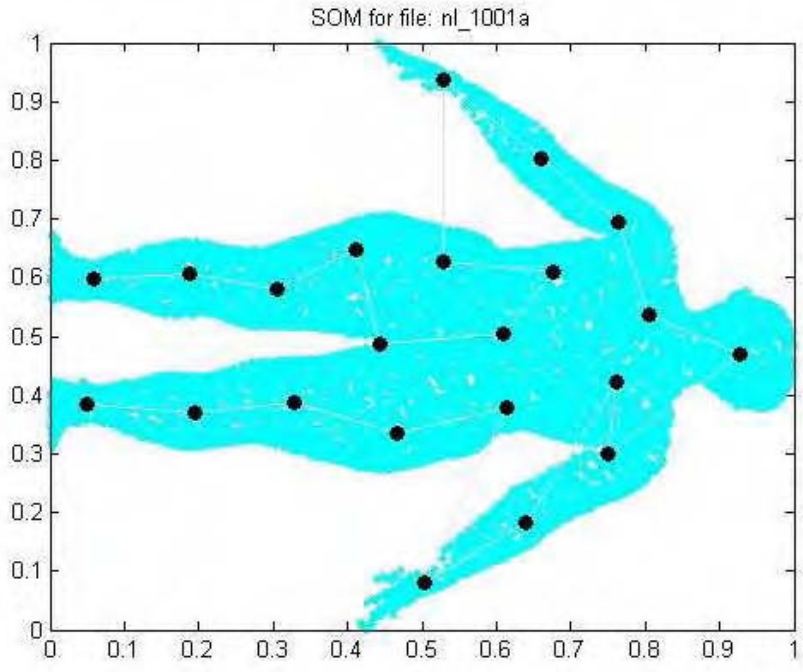


Figure 33. Representative result of the SOM portion of the marker estimation algorithm.

3.3.2 Alpha Hull.

Geometry computation is accomplished through the use of the alpha hull. As described in Sec. 2.3.4.3, the choice of disk radius has a significant impact on the resultant shape. So, in selecting an appropriate radius is an area where trade offs are weighed to provide the best results. If too large of a radius is selected, then the resultant shape becomes a convex hull as in Fig. 34; whereas, if the radius is too small, areas of missing data may cause discontinuities and gaps in the shape. The appropriate value for the alpha hull radius was found using empirical methods. The output of the alpha hull portion of the algorithm is shown in Fig. 35.

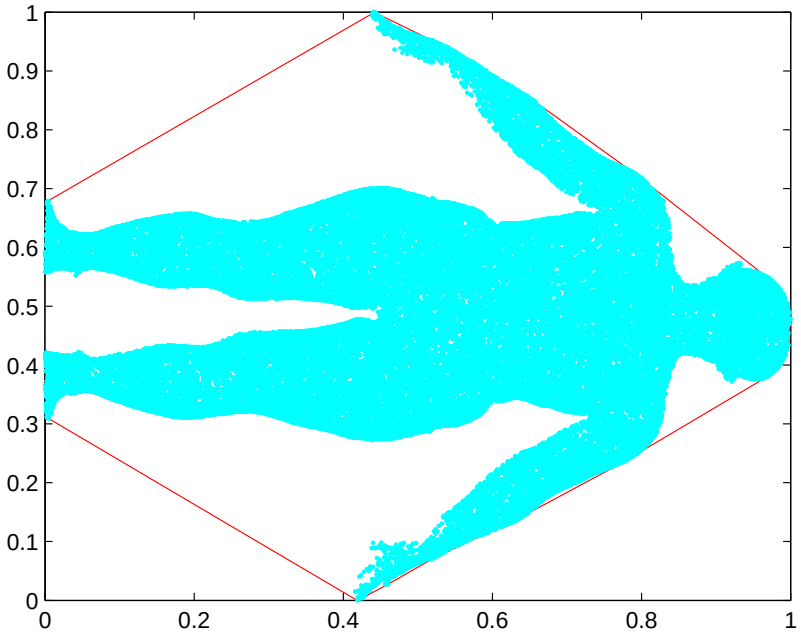


Figure 34. Results of the alpha shape algorithm when $\alpha = 0.75$. The α is too large to provide fidelity to the outline shape of the subject.

The resulting alpha hull serves as the initial step in pruning the SOM lattice. The nodes are compared to the alpha hull, and if the node falls outside of the alpha hull, that node is eliminated from the marker set. Further, the alpha hull provides connectivity information based on testing whether the midpoint of each lattice connection falls within the alpha hull. If the midpoint falls outside the alpha hull, that connection is discarded. A final reason

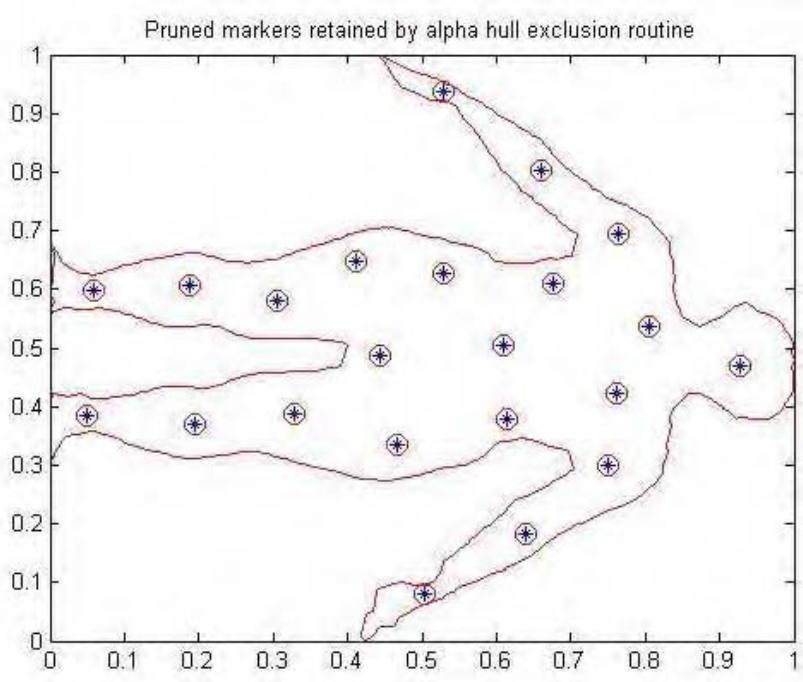


Figure 35. Results of the alpha hull algorithm. The asterisks indicate the location of a lattice node from the SOM. Circles around the asterisk indicate nodes that lie inside the alpha hull based upon a positive result from a point-in-polygon test.

for employing an alpha hull is that it is the initial step in developing a skeleton for the projected data set.

3.3.3 Voronoi Skeleton.

The medial axis, or pruned Voronoi tessellation [22], is used to find the optimized skeleton based on the alpha hull. The resultant skeleton provides a refinement to the centralized markers: ankle, knee, wrist, elbow and shoulder. A typical output from a basic Voronoi skeletonization is shown in Fig. 36. The results are first pruned based on what is inside the alpha hull, as shown in Fig. 37. Since the general shape of the desired skeleton is known, further refinement is done by pruning the Voronoi diagram based on removing lines that have a slope inconsistent with $slope \approx 0$ as shown in Fig. 38. In practice, the pruning portion of the algorithm is not as successful as desired and does not reliably provide the skeleton information desired which is the mid line of the alpha hull for specific areas: the legs, arms, and head. These are the areas needed to refine the anthropometric markers to

the final location. With this goal in mind, the algorithm simplifies the Voronoi process by scanning through all of the points along the x-axis and noting the points of intersection with the alpha hull. Each pair of intersections is averaged to find the midpoint which then becomes a constituent component in the skeletal object. An example of the output from the described revised medial axis algorithm is shown in Fig. 39. The lines connecting the legs to the torso are not computed since they do not follow the skeletal structure. The simplified method of computing the medial axis is used since the Voronoi diagram fails to provide desirable information around the hips; therefore, using the more complex method does not improve the results.

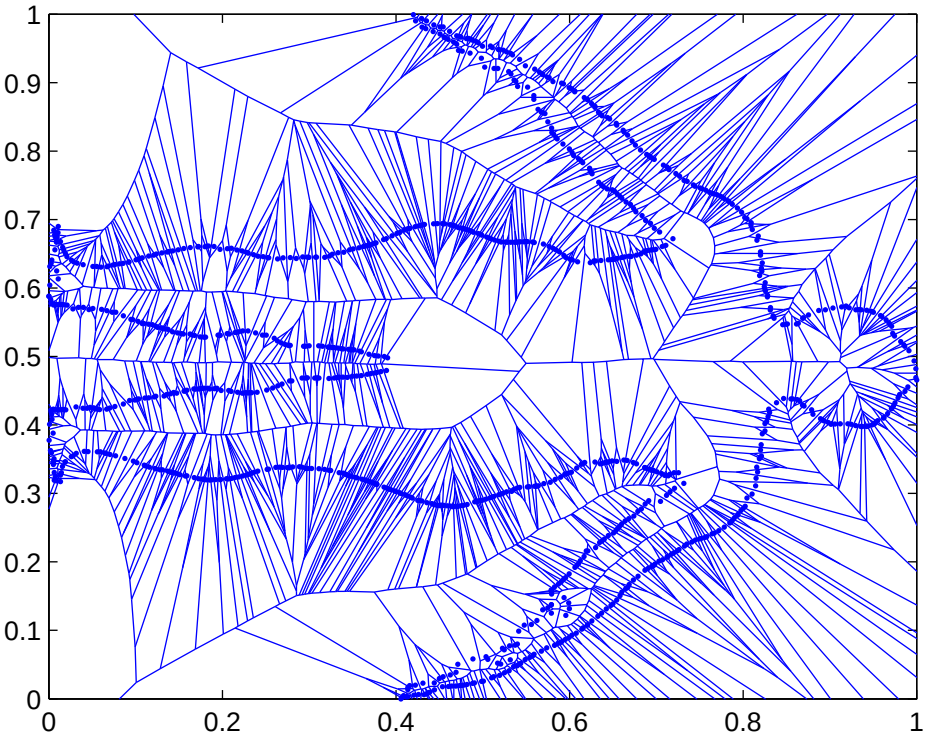


Figure 36. A Voronoi diagram from the alpha hull obtained from the projected 3-D point cloud of a subject from [72].

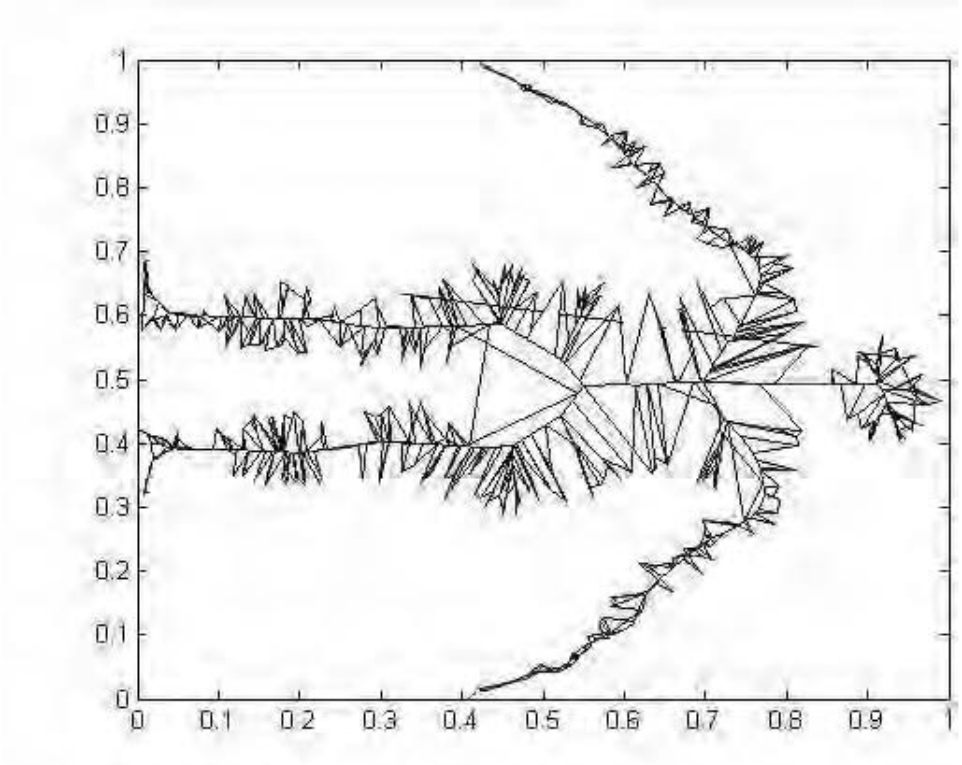


Figure 37. The initial step in pruning the Voronoi diagram by removing the lines not contained within the alpha hull.

3.3.4 Incorporating Anthropometric Information into the Algorithm.

Due to the potential for misalignments of the SOM neurons as estimates for the anthropometric marker locations, we exploit the structure of the human body to derive the estimated markers. The next portion of the algorithm utilizes anthropometric information to refine the final marker locations. Part of this component of the algorithm uses basic anthropometric information to assign locations and labels to the markers, e.g. the next connected marker away from the ankle will be the knee, etc. This provides the labels for the markers. The locations get refined based on empirically determined results. Typical results for this stage of the algorithm are presented in Fig. 40.

Based on the converged SOM map and the resulting output of the alpha hull, the anthropometric pruning refines the location of the final markers. There are a number of checks that need to be considered based upon the available information, such as (a) whether the platform is included in the scan necessitating the incorporation of internal alpha hulls, (b) whether the 3-D body scan is corrupt or missing parts, and (c) whether there are

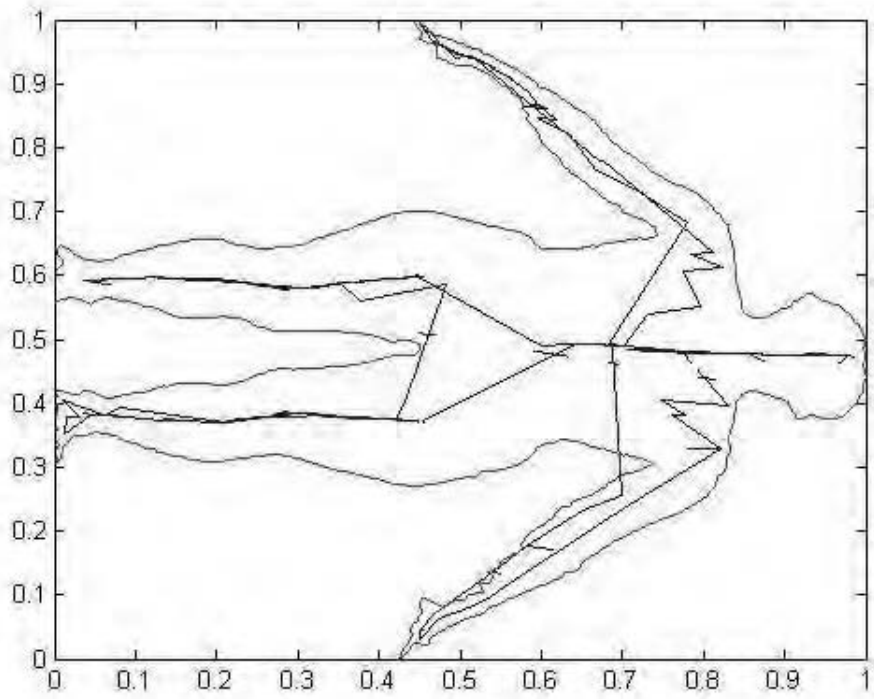


Figure 38. The skeleton from the Voronoi diagram results derived from the alpha hull of the projected point cloud for a subject from [72].

erroneously excluded SOM lattice nodes. If the scanned data are too sparse to fully connect the subject, the algorithm makes some assumptions to fully flesh out the marker set. One assumption made is that if one ankle marker is found, there will be a corresponding marker in the other leg. Similarly, if the scan did not completely capture parts of the arm, then the algorithm assumes that there is a wrist or elbow marker in an expected location, e.g. near the end of the arm or between the wrist and shoulder markers. There are also instances when the alpha hull excludes an SOM lattice node that appears to be inside the resulting polygon. Once again, on these occasions, the anthropometric algorithm places a marker in the missing location corresponding to the one found in the other appendage. (The omission of some elements inside the alpha shape is likely an error in the specific MATLAB® implementation. However, adapting to the MATLAB® issue is useful in the case when there are missing/incomplete portions of the scan.)

Based on the shape resulting from the alpha hull, anthropometric pruning may perform differently. If the alpha hull is of the expected format where the scan is complete and there is

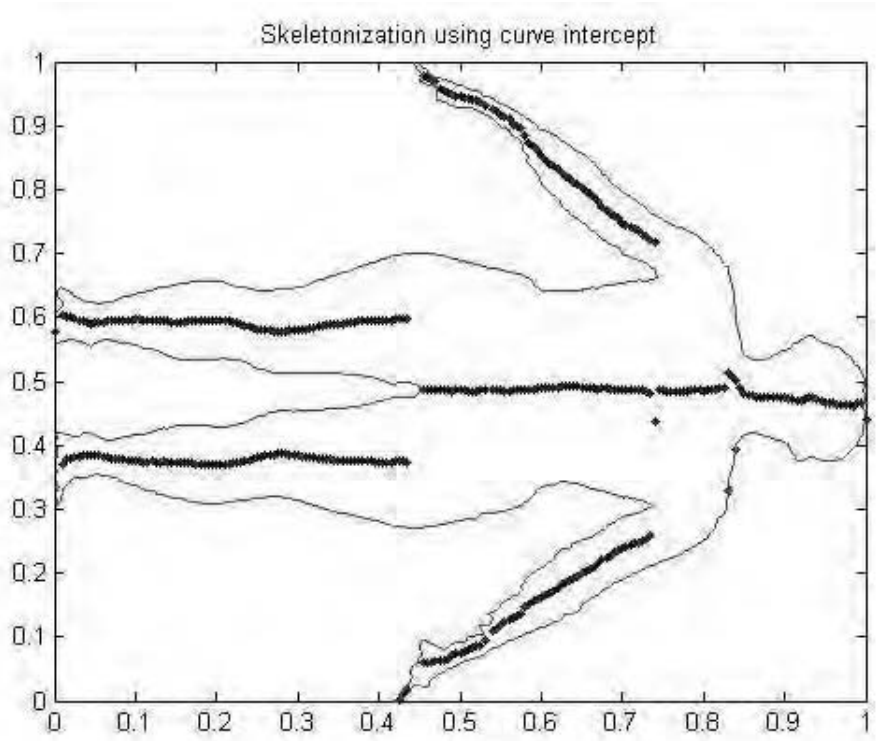


Figure 39. The medial axis results derived from the alpha hull of the projected point cloud for a subject from [72].

distinct separation between the arms and the torso and between both legs as that presented in Fig. 41, then anthropometric pruning relies on the Voronoi based skeletal results to center the markers in the appendages. If the result of the alpha hull is similar to that in Fig. 42 where the definition between the legs is not captured, then the algorithm looks to incorporate an internal alpha hull piece that removes the section between the subject's legs that does not correspond to data. If the result is not available, then the algorithm relies on the initial placement of the nodes from the SOM mapping.

For the knee and hip markers, the final marker location is determined by averaging the location of the two lattice nodes that fall closest to the anticipated marker locations based on the golden section results in Section 2.3.6.1 [63]. As noted in Section 2.3.6.1, the experimental results disproved the validity of basing body measurements solely on the golden segment principle. However, general estimates for where a marker may lie can be made. In this case, in the normalized space in which the algorithm works, knees are expected to be found around the $x = 0.25$ mark and hips around the $x = 0.45$ lines. These values

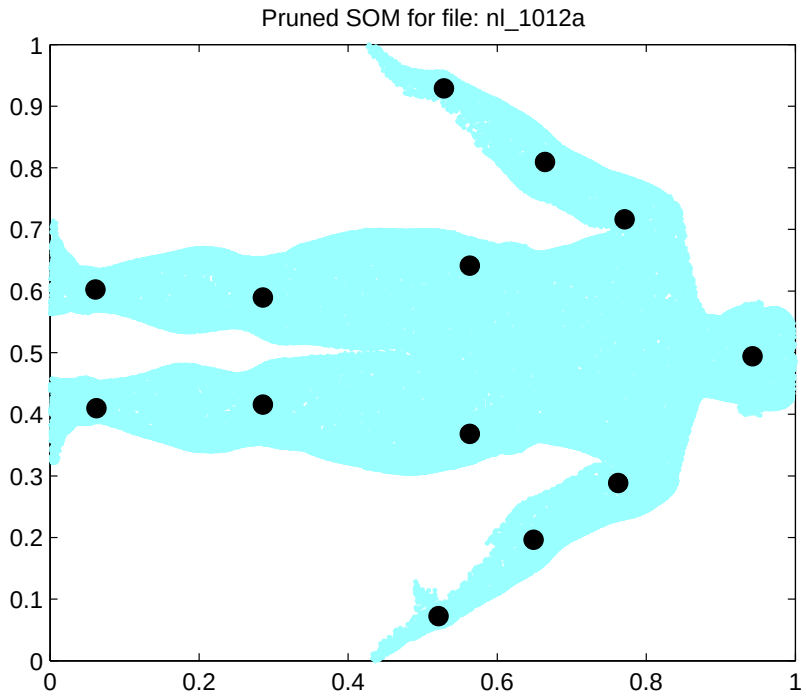


Figure 40. Results of the pruning and anthropometric adjustment portion of the algorithm.

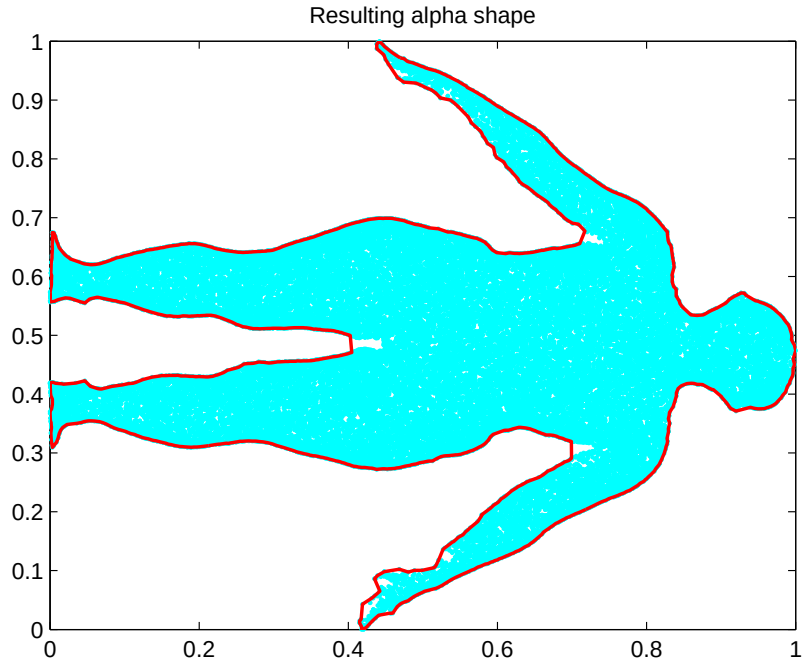


Figure 41. The resulting alpha hull with $\alpha = 0.025$.

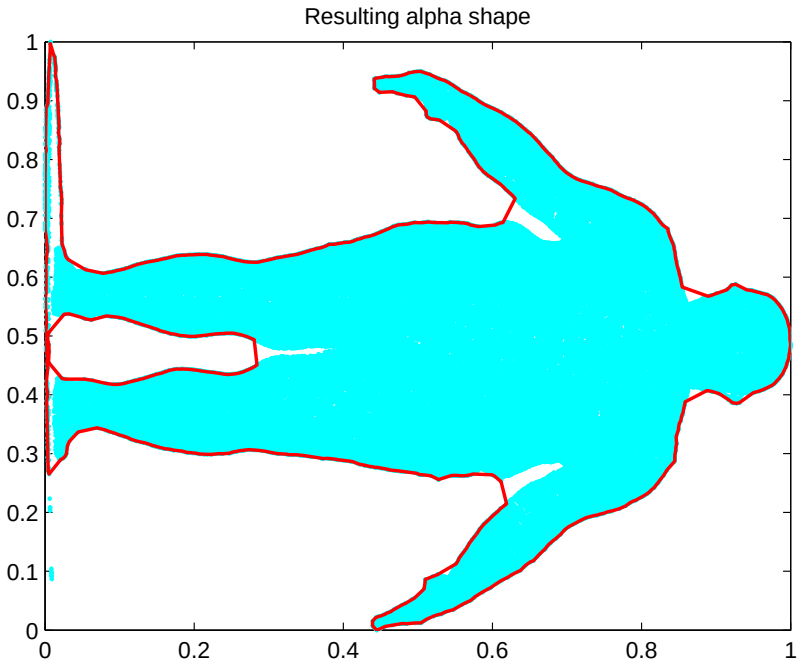


Figure 42. The resulting alpha hull with too much of the platform showing in the scan.

do not directly factor into the calculation of the stated markers, but only serve as the initialization point for searching for the two nearest neurons to use. Based on the number of nodes used in the SOM portion of the algorithm, the location of the markers along the skeletal line for the knees is found by averaging the two nearest nodes to the $x = 0.25$ line which results in finding x_{knee}. The process is done for both knees. The x-axis location of the hips is found by averaging the two nodes closest to the $x = 0.45$ line. This results in x_{hip} for one of the hips. This process is repeated for the other hip as well which results in a pair of x-axis locations for both the knees and hips.

Having found the ankle and knee markers, the hip marker refinement process begins by finding the intersection of the line formed by the ankle-knee markers and the vertical line associated with the averaged hip marker location. This intersection point becomes the y-axis location for the hip. This process is repeated for the other leg.

The anthropometric refinement process for the arms is minimal. The results from the SOM process provide three markers located in each arm as seen in Fig. 43. The results from the skeletal representation is used to center the marker along the width of the arm.

Occasionally, the results of the SOM place two nodes in the head. In these instances, the algorithm averages the locations of the two markers and projects the results onto the skeleton.

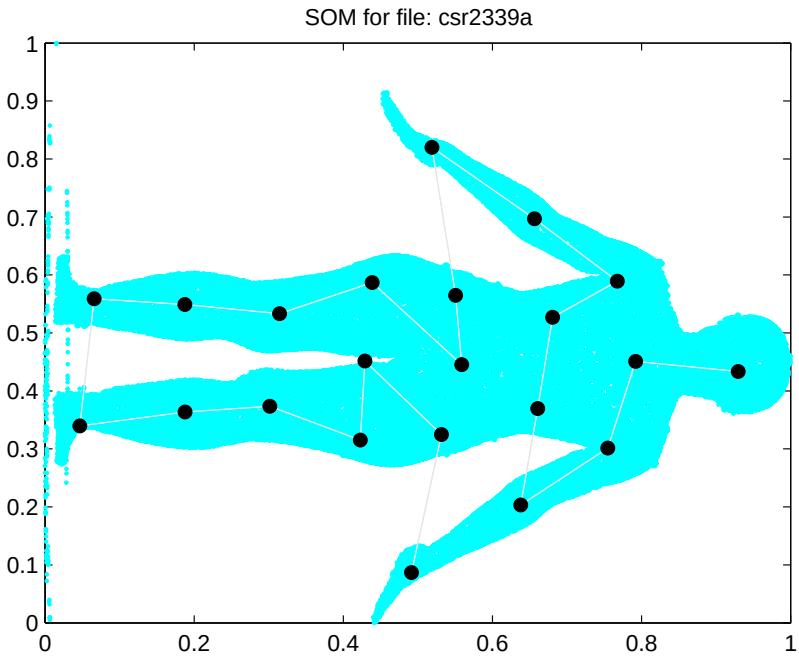


Figure 43. An example of the typical lattice node locations in the arms for the 1-D SOM.

3.3.5 Post-processing.

The resulting set of markers is rescaled using the retained scaling factors from the original preprocessing step. The feature set used in the classification phase is then calculated based on calculating the distance between the appropriate markers to obtain the body segment lengths.

3.4 Summary

This chapter described the various components of the marker estimation algorithm and the methodology used in selecting those components. In the next chapter the specific details of the implemented algorithm are provided and the results from the classification are presented and analyzed.

4. Experimental Results and Analyses

This chapter documents the simulation process, the parameter choices made for each phase, and the results of the point cloud segmentation algorithm developed in this thesis and depicted in Fig. 31b. Additionally, the results of the gender classification are presented. The chapter begins by discussing of each part of the segmentation algorithm and describes and justifies parameter choices followed by the results obtained through classification. The final section of this chapter presents an analysis of the results to include computational efficiency comparing those methods discussed in the literature.

4.1 Data

The 3-D LADAR point clouds come from the CAESAR database [72]. The point cloud is generated by the body scanner's four scanning heads which provide a 360° scan of the subject as can be seen in Fig. 44. Each subject's point cloud consists of approximately 250,000 data points. The overall CAESAR database consists of 4,426 subjects. However, due to missing data and markers, a small subset of those scans are not used in this research.

Table 4 provides the basic demographic information for the CAESAR database in total including the number of subjects from each country and a breakout of the gender distribution. Tables 5, 6, and 7 provide the make up of each country's database separately.

Table 4. Demographics for the test subjects.

Scan Location	Males	Females	Totals	Percentages
North America	1120	1255	2375	53.7%
Italy	410	386	796	18.0%
Netherlands	564	691	1255	28.3%
Totals	2094	2332	4426	
Percentages	47.3%	52.7%		

A total of 3825 subjects are used in the classification process to evaluate the segmentation algorithm. The distribution of the test set is approximately equally distributed between male and female subjects.

Figure 44. The type of scanner used to obtain 3-D body scans in the CAESAR database.

Table 5. Racial demographics for the North American subjects.

Race	Males	Females	Totals	Percentage
White	867	957	1824	76.8%
Black	116	147	263	11.1%
Other	137	151	288	12.1%
Totals	1120	1255	2375	

4.2 Experimentation Process

4.2.1 Monte Carlo Simulations.

A Monte Carlo simulation on each subject is conducted with the segmentation algorithm. This is done for 100 runs on 40 of the subjects to determine the distribution of the calculated markers. An example of the distribution of the calculated markers for a random subject

Table 6. Nationalities of subjects from the Netherlands database.

Nationality	Males	Females	Totals	Percentage
Dutch	480	544	1024	81.6%
Other	84	147	231	18.4%
Totals	564	691	1255	

Table 7. Nationalities of subjects from the Italian database.

Nationality	Males	Females	Totals	Percentage
Italian	388	376	764	96.0%
Other	22	10	32	4.0%
Totals	410	386	796	

is shown in Fig. 45. Histograms for a representative sample of marker points is presented

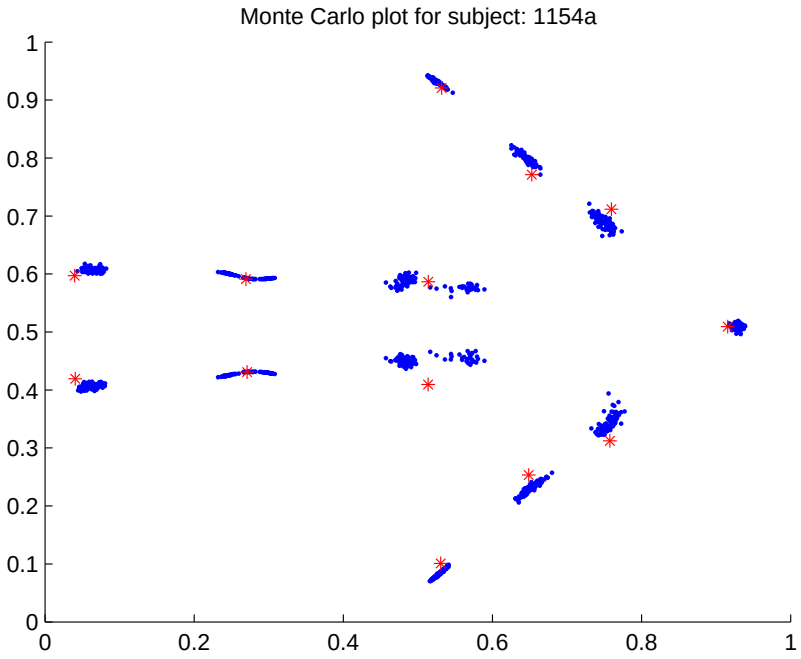


Figure 45. An example of the distribution of calculated markers from the Monte Carlo simulation for a random subject.

below in Figs. 46 and 47. Laying the 3-D histograms over the projection of the point cloud data is shown in Fig. 48. Due to processing time considerations, it is necessary to reduce the number of simulations. This is possible if the data are approximately normally distributed. In that case 10 simulations would suffice. The subject of multivariate normality is explored in further detail in Section 4.4.

4.2.2 Classification Process.

For the classification process, three approaches are used. When a separate test set is available, it is used. If a separate test set is not available, then the training data set is used

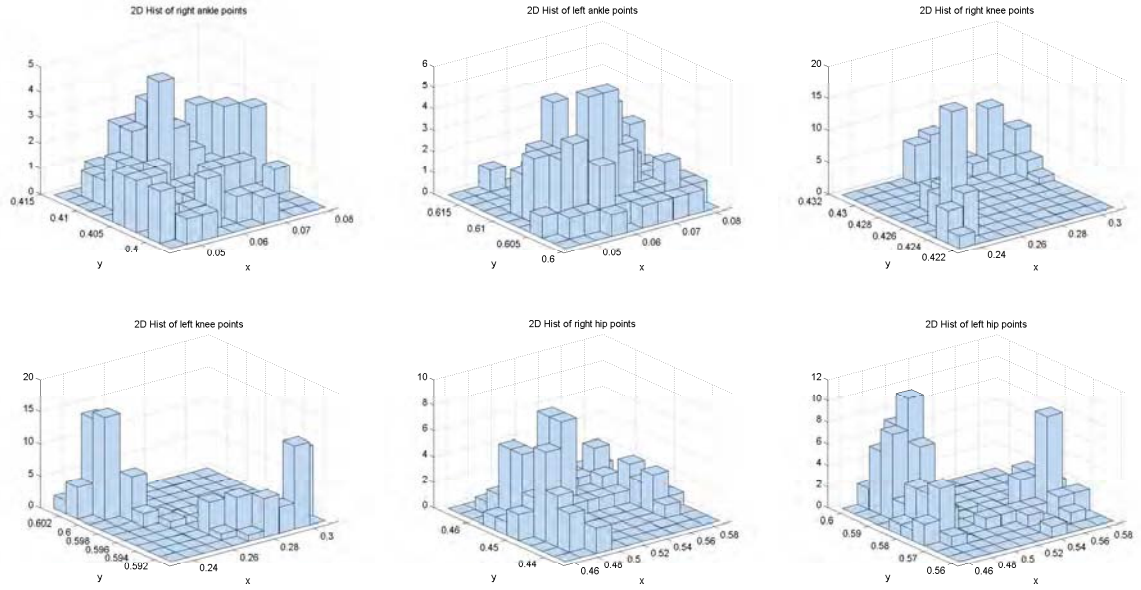


Figure 46. Part 1 Histograms for anthropometric marker locations based on 100 Monte Carlo simulations.

as both the training and the testing data through the use of resubstitution and K-fold cross validation. In resubstitution, the entire data set is used for training and testing. In K-fold cross-validation, the training data is divided into K approximately equally sized groupings. One of the subdivided groups is withheld from the training step and is used as the test set. The other $K-1$ groups are used to train the classifier. This process is repeated for each of the other $K-1$ groups. This provides an estimate for the true generalization error.

Classification is done for four separate cases in which the test data and truth data are comprised of the following data sets, respectively: Case I, truth markers versus truth markers; Case II, estimated markers versus truth markers; Case III, Monte Carlo mean markers versus Monte Carlo mean markers; and Case IV, estimated markers versus Monte Carlo mean markers. In Cases I and III, resubstitution and K-fold cross validation are used in the classification. In Cases II and IV, separate test data are available and used in the classification process.

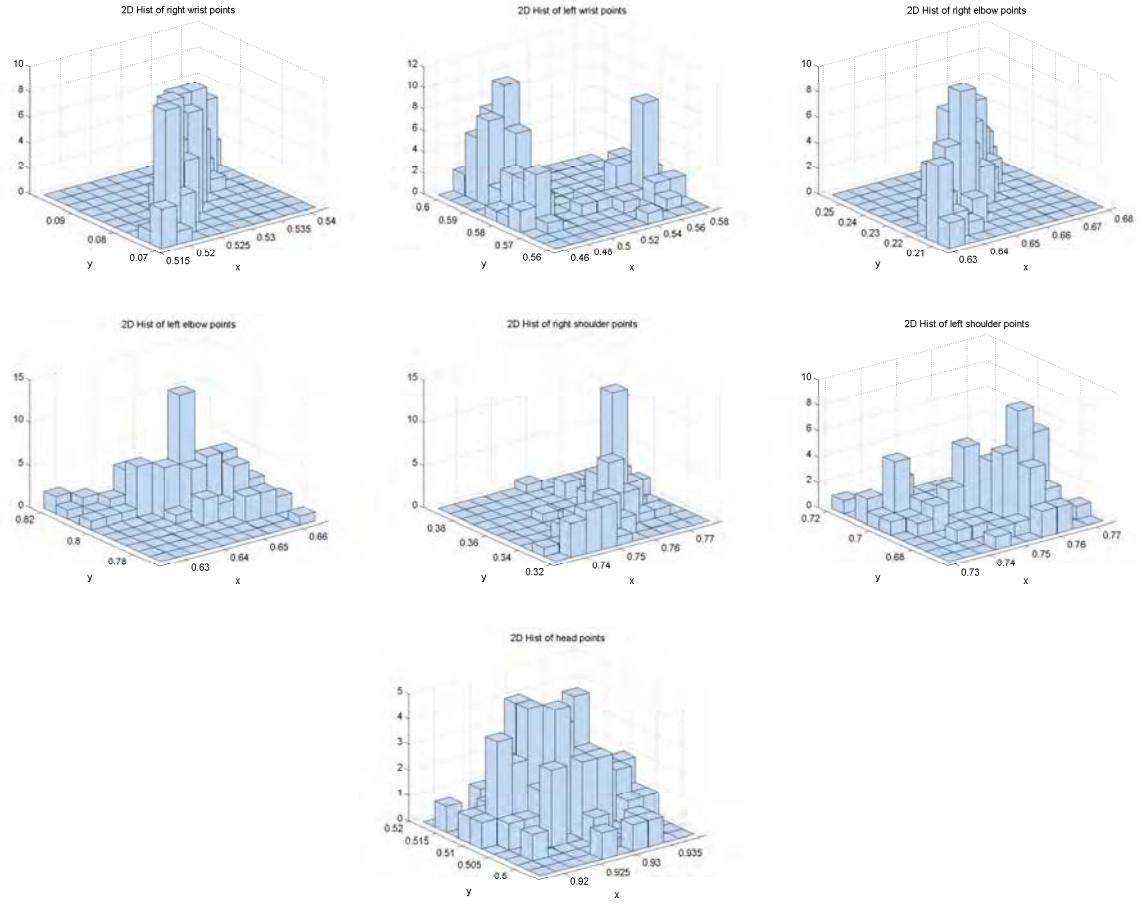


Figure 47. Part 2 Histograms for anthropometric marker locations based on 100 Monte Carlo simulations.

4.2.3 Classification Evaluation Methods.

The first step in reporting the classification results is to recreate the results from the sponsor's paper on gender classification [27]. This is discussed in Section 2.4.1. Using the entire 16-measurement feature set shown in Table 1, the developed classifier duplicated the 98.5% classification accuracy for the Netherlands database and 99.5% classification accuracy for the North American and Italian databases through the use of resubstitution. When the feature set used for classification is reduced to the distance measures listed in Table 3, the classification accuracy dropped to 88.7% for the truth data when performing resubstitution.

In each of the four cases described in Section 4.2.2, the performance of the classification process is reported for the entire CAESAR database as well as for the database subsets

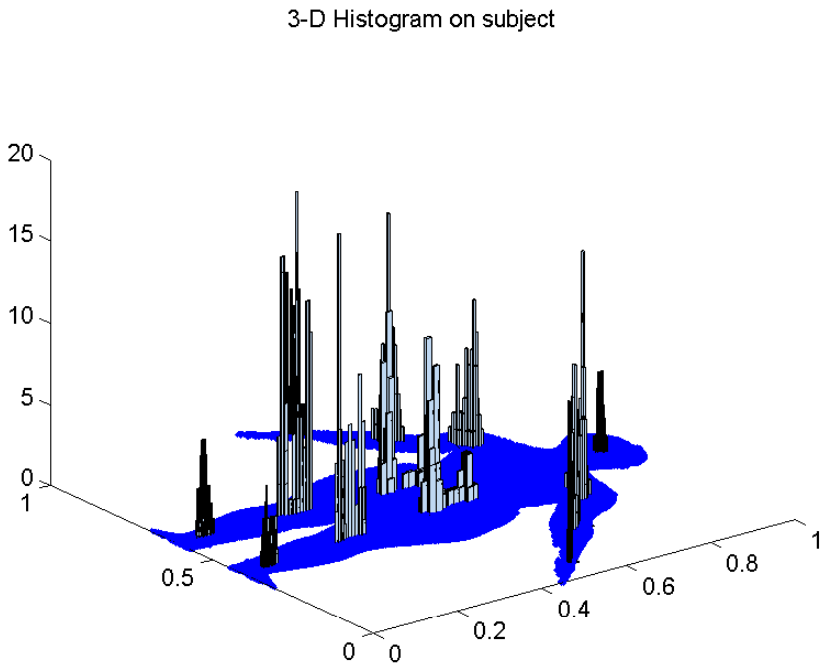


Figure 48. An example of the 3-D histograms located at their respective marker locations.

consisting of the scans for North America, Italy, and the Netherlands. For each of the subsets, the testing and training data are constrained to come from that subset alone, e.g. the North American truth markers are classified versus North American truth markers.

4.3 Segmentation Algorithm Parameters

Many of the marker estimation algorithm components depicted in Fig. 31b rely on specific parameter choices. Those parameters and results are discussed here.

4.3.1 Principal Component Analysis.

The 3-D body scan data is processed through the principal component analysis (PCA) algorithm retaining the first two principal components. At this stage, the data are also scaled. The results of these two steps are shown in Fig. 32.

4.3.2 Self-Organizing Map.

When using the Self-Organizing Map (SOM), there are several parameters that need to be determined. Namely, the dimension of the lattice, the number of nodes in the lattice, the neighborhood function, and the learning rate are all available to be adjusted based on the application.

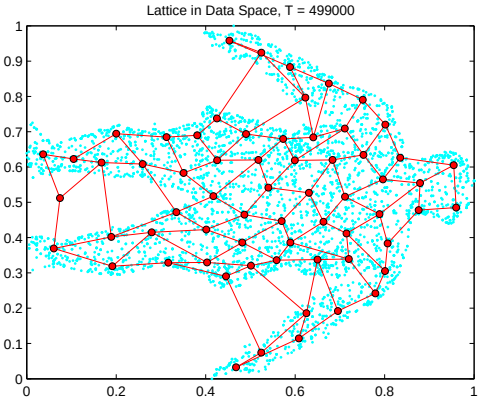
4.3.2.1 Lattice Dimension.

Initial exploration of both 1-D and 2-D lattices were explored early in the research. Some examples of the initial explorations using a 2-D lattice on the data in 2-D PCA space and mapping the lattice to data space results in the plot shown in Fig. 49a. The SOM is comprised of an 8 x 8 lattice of neurons. Looking at different lattice sizes results in the mappings presented in Fig. 49b for a 5 x 5 lattice and Fig. 49c for a 10 x 10 lattice. Applying a 1-D lattice to the 2-D data results in the mapping shown in Fig. 49d for a 20 neuron lattice. A 30 neuron lattice is presented in Fig. 49e and a 40 neuron lattice is presented in Fig. 49f. The 1-D lattice is selected based upon the desired end results. Since the goal is to determine anthropometric markers that align to a human skeleton and much of the skeletal framework of interest is a linear entity, the 1-D lattice is used.

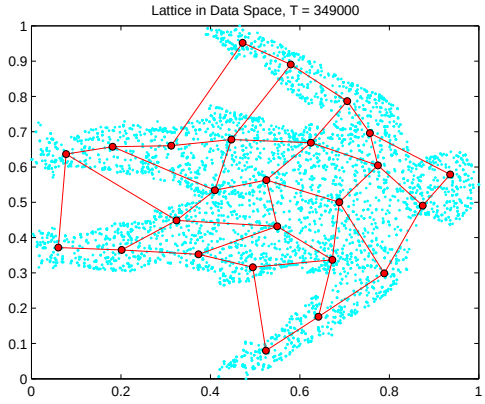
4.3.2.2 Number of Neurons.

The number of lattice nodes is initially explored empirically with a routine utilized to determine the number of nodes that best fit the data. The result of sweeping through a range of nodes and calculating the average topographic error [39, 85] is shown in Fig. 50. The quantization and topographic errors are a significant factor in determining the appropriate number of nodes selected since the desire is to minimize *both* errors.

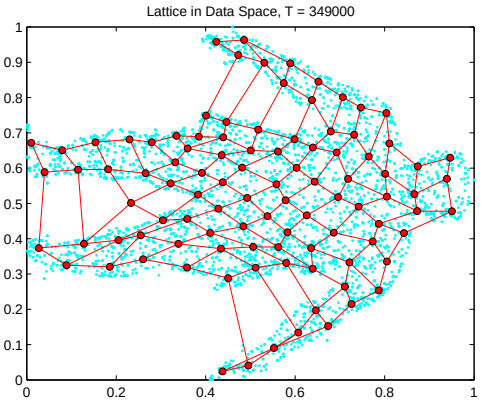
A widely used metric to determine the quality of the resulting SOM is the quantization error. This error measures the average distance between each data vector and its best



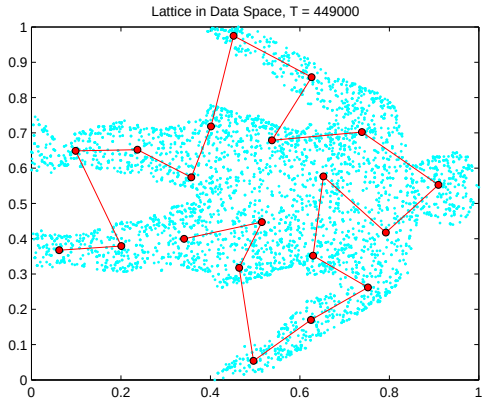
(a) 8 x 8 lattice SOM in projected data space.



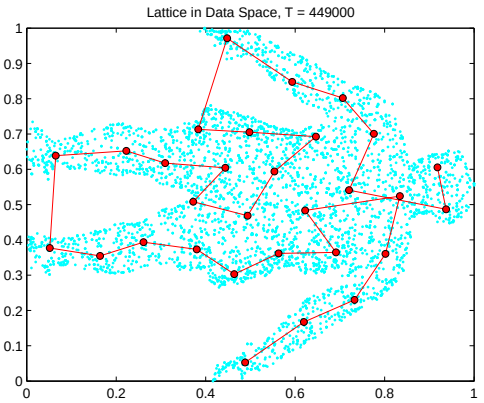
(b) 5 x 5 lattice SOM in projected data space.



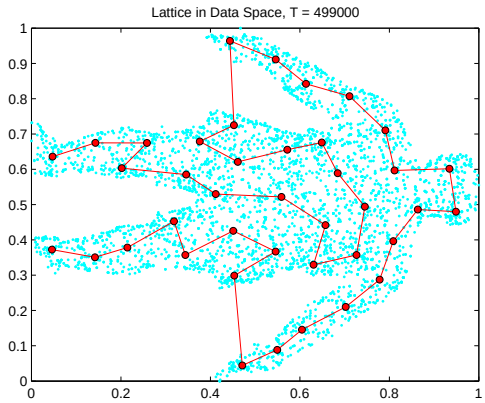
(c) 10 x 10 lattice SOM in projected data space.



(d) 20 neuron 1-D lattice SOM in projected data space.



(e) 30 neuron 1-D lattice SOM in projected data space.



(f) 40 neuron 1-D lattice SOM in projected data space.

Figure 49. SOMs of various lattice sizes shown in the projected data space.

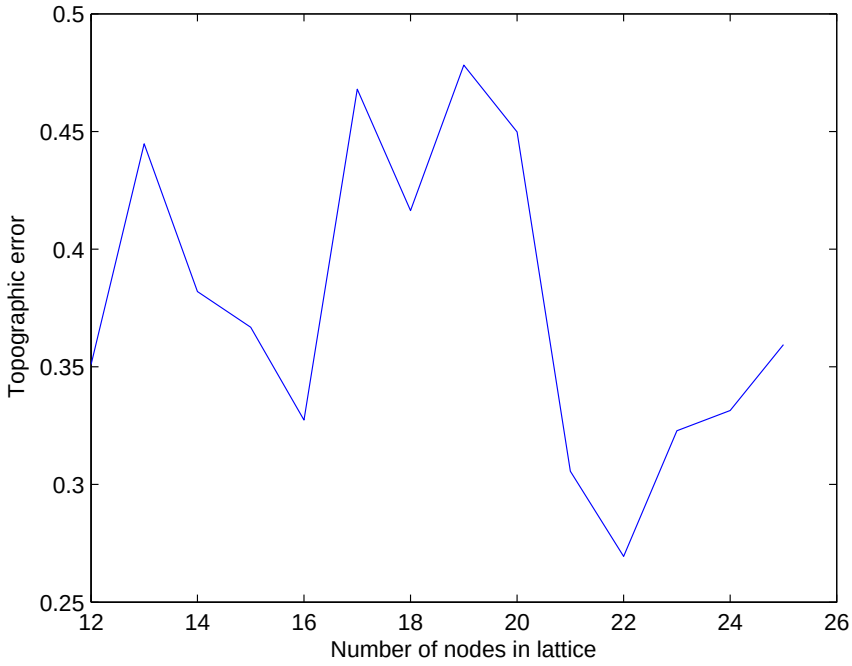


Figure 50. The topographic error for a 1-D lattice SOM in data space.

matching unit (BMU) [46]. The quantization error is calculated as:

$$\varepsilon_q = \frac{1}{N}\sum_{i=1}^{N} ||\vec{x}_i - m_{\vec{x}_i}|| \tag{34}$$

where N is the number of data vectors and $m_{\vec{x}_i}$ is the BMU of the corresponding $\vec{x}_i$. This error evaluates how well the neural map fits the data. Therefore, the smaller the quantization error, the closer the data vectors are to the prototype vectors [85]. Quantization error decreases as the number of lattice nodes increase.

An error measure used to quantify the topological preservation of the map is the topographic error. For a given $\vec{x}_k$, the SOM finds the BMU (w_i) and the second-BMU (w_j). Input vectors between $\vec{x}_k$ and w_i are mapped to w_i or w_j. The mapping is called locally continuous if the associated nodes for w_i and w_j are adjacent. If they are not adjacent, then there is a discontinuity which is captured as a local topographic error. The overall topographic error is defined as the proportion of local topographic errors that exist in the

map and is formalized as:

$$\varepsilon_t = \frac{1}{N}\sum_{k=1}^{N} u(\vec{x}_k) \text{ where } u(\vec{x}_k) = \begin{cases} 1, & \text{if best and second BMUs are non-adjacent} \\ 0, & \text{otherwise} \end{cases} \tag{35}$$

A noted failing, or feature, of this measure is that it does not account for the closeness of the local topographic error. It does not matter if two data vectors that are close to one another in the data space get mis-mapped to nodes that are close, but not adjacent, or to the extreme opposite ends of the map [43].

4.3.2.3 Neighborhood Function.

The utilized neighborhood function is a time-decaying Gaussian (exponential) of the form:

$$h_{j,i(x)}(n) = exp\left(-\frac{d_{j,i}^2}{2\sigma(n)^2}\right) \tag{36}$$

where $d_{j,i}$ is the distance between the winning neuron i and the excited neuron j (in lattice units, not weight vectors).

The size of the neighborhood shrinks over time so that the width (σ) takes the form:

$$\sigma(n) = \sigma_0 exp\left(-\frac{n}{\tau_1}\right) \quad n = 0, 1, 2, ... \tag{37}$$

where σ_0 is the initial value for σ, τ_1 is a time constant, and n are discrete time units.

4.3.2.4 Learning Rate.

The learning rate is a stepped, decaying exponential as shown in Fig. 51. The learn rate is held constant for an initial period of time and then decayed exponentially. A common form for the learning rate is [35]:

$$\eta(n) = \eta_0 \exp\left(-\frac{n}{\tau_2}\right) \tag{38}$$

where η_0 is the initial value, τ_2 is another time constant, and n are discrete time units. According to [35], the learning rate should begin with a value close to 0.1, decrease gradually from this initial point, but remain above 0.01. Based on these guidelines, it was found that in practice, learning rates of the form prescribed in Eq. 38 decayed too rapidly, so the learning rate used in the marker estimation algorithm is:

$$\eta(n) = \eta_0 n^{-0.1} \tag{39}$$

where $\eta_0 = 0.1$.

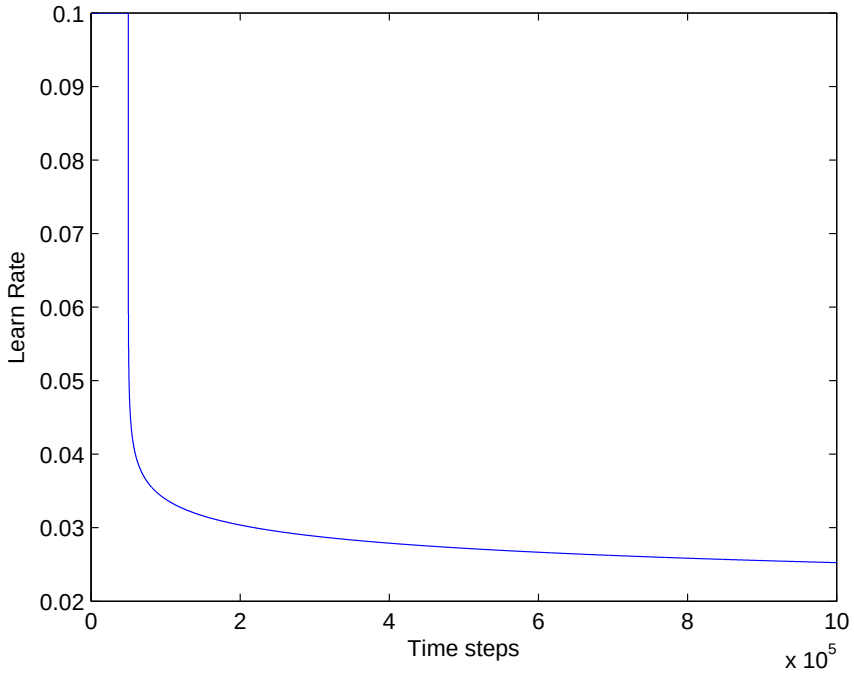


Figure 51. The learning rate used in the SOM training.

4.3.3 Alpha Hull.

A representative alpha hull result, when applied to projected data, is shown in Fig. 52. Here, a radius of $\alpha = 0.025$ is used. The alpha hull combined with the converged SOM lattice nodes is shown in Fig. 53. In this example, one node located near the crotch of the subject is excluded by the alpha hull, as indicated by the red asterisk without a blue circle

around it. Based on how some of the scans were conducted, there is a wide variation in whether the platform is included in the scanned volume or not. In the North American scans, a large portion of the scanner's platform appears in the scans. This presents a challenge when generating the alpha hull (Fig. 54). If the radius of the alpha hull disk is too large and the platform is in the scan, we get results similar to those shown in Fig. 55. Table 8 provides a listing of the measured platform offsets at the various scanning locations. For some of the data sets, it is necessary to remove the platform offset from the data prior to projecting it. This alleviates the inconvenience associated with the platform.

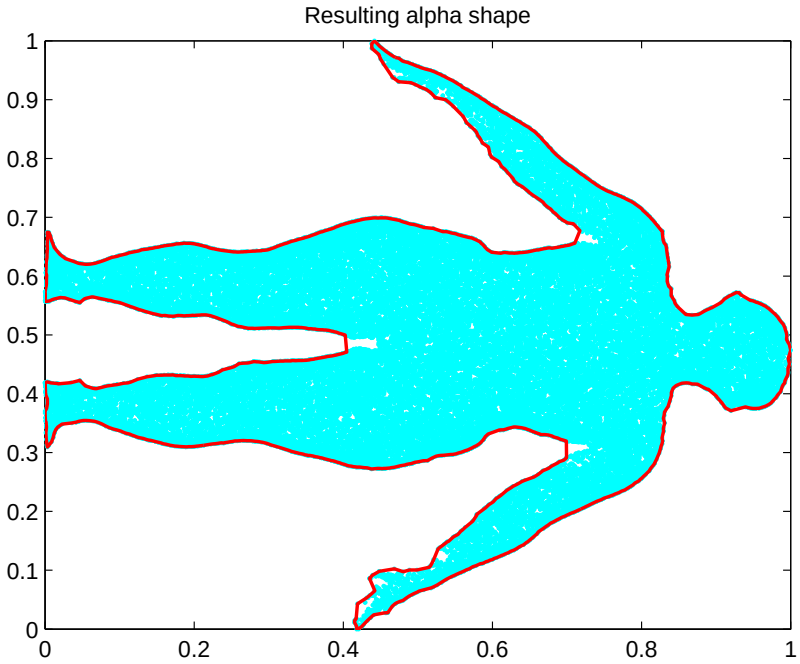


Figure 52. The resulting alpha hull with $\alpha = 0.025$.

4.3.4 Monte Carlo Simulation Numbers.

A Monte Carlo simulation on each subject is performed using the segmentation algorithm. Due to processing time considerations, as mentioned in Section 4.2.1, a Monte Carlo simulation consists of 10 runs of the marker estimation algorithm on each subject. The 10 runs are averaged to generate a set of mean marker locations for use as training data in the classification process.

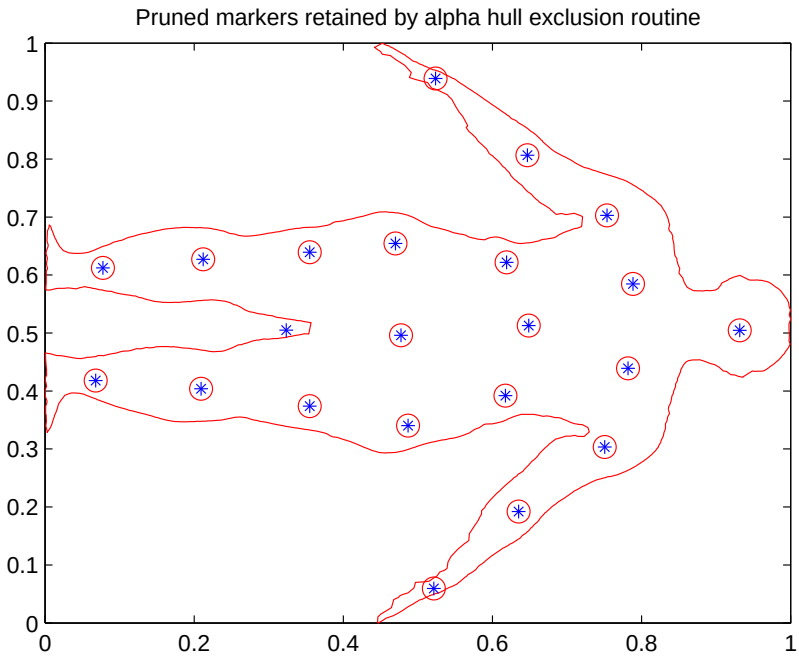


Figure 53. The set of markers retained by the alpha hull routine. SOM nodes are denoted by an asterisk. Retained markers are denoted by a circle around the retained node.

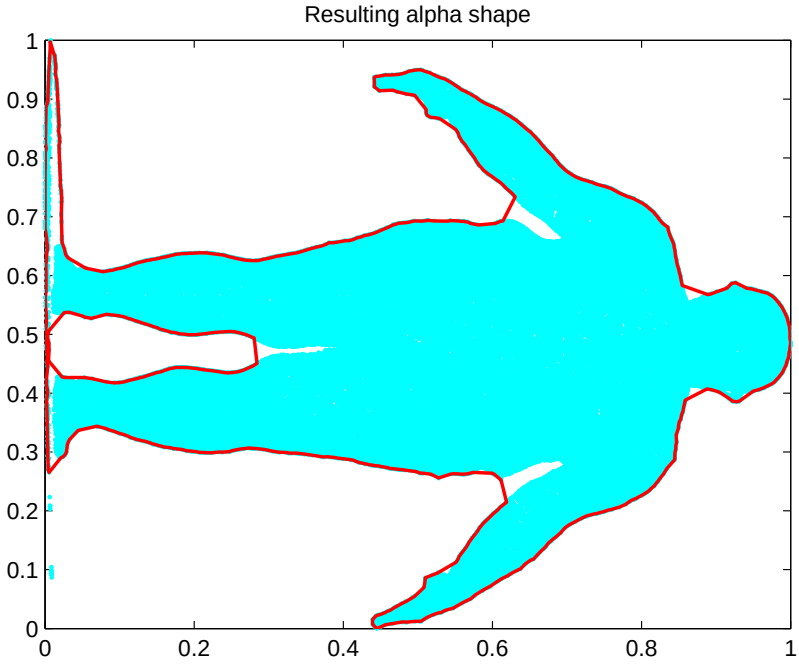


Figure 54. The resulting alpha hull with too much of the platform showing in the scan.

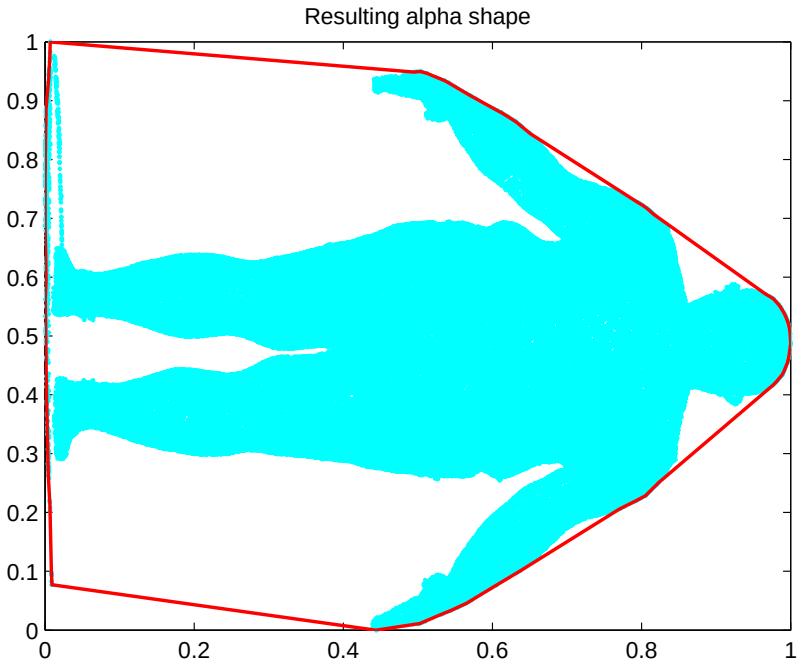


Figure 55. The resulting alpha hull with too large of an α and too much of the platform showing in the scan.

4.4 Marker Estimation Error

In addition to evaluating the functional performance of the marker estimation algorithm through gender classification, the estimated marker locations are examined for multivariate normality, and their distance from the mean estimated marker and truth marker locations is calculated using the Mahalanobis distance.

4.4.1 Multivariate Normality.

In order to justify the reduction in the number of runs in the Monte Carlo simulation, the distribution of markers is tested for normality using various methods. An initial empirical graphical method served to indicate the likelihood of normality. Based on positive results from the empirical analysis, further tests are performed to confirm normality, specifically, the Mardia's, Royston's, and Henze-Zirkler's tests for normality [58, 59].

Table 8. Platform offsets for the various data collection sites.

Data Collection Site	Subject Numbers	Platform Offset (mm)
1. Los Angeles, CA	0001-0321	1000
2. Detroit, MI	0324-0671	992
3. Ames, IA	1031-1173	998
4. WPAFB1, OH	1174-1261	994
5. Greensboro, NC	1262-1346	998
6. Marlton, NJ	1350-1480	1000
7. Ottawa, Ontario, Canada	1481-1713	1000
8. Minneapolis, MN	1715-2119	998
9. Houston, TX	2120-2343	998
10. Portland, OR	2344-2592	998
11. WPAFB2, OH	2593-2618	996
12. San Francisco, CA	2619-2788	998
13. Atlanta, GA	2789-3028	998
14. Genoa1, Italy	4000-4394	1000
15. Genoa2, Italy	4395-4800	1000
16. Soesterberg, Netherlands	11001-17079	993.1

4.4.1.1 Mardia Test.

If a set of variables is multivariate normally distributed, then each variable must be normally distributed. However, the reverse is not necessarily true as the set of variables may not be multivariate normally distributed just because all of the individual variables are normally distributed. Hence, testing each variable only for univariate normality is not sufficient. Mardia proposed tests of multivariate normality based on sample measures of multivariate skewness and kurtosis [54].

For a random variable $\mathcal{X} = (x_1, ..., x_n)$ from a p-variate distribution with a sample mean, $\bar{x}$, and sample covariance matrix, S, Mardia defines the p-variate skewness and kurtosis statistics by the following

$$b_{1,p} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[(x_j - \bar{x})^T S^{-1} (x_i - \bar{x}) \right]^3, \tag{40}$$

$$b_{2,p} = \frac{1}{n} \sum_{i=1}^{n} \left[(x_i - \bar{x})^T S^{-1} (x_j - \bar{x}) \right]^2. \tag{41}$$

Given that the data is multivariate-normally distributed, the expected value of the

multivariate skewness statistic is given by

$$\mathbb{E}[b_{1,p}] = \frac{p(p+2)[(n+1)(p+1)-6]}{(n+1)(n+3)}, \tag{42}$$

and the expected value of the multivariate kurtosis statistic is

$$\mathbb{E}[b_{2,p}] = \frac{p(p+2)(n-1)}{(n+1)}, \tag{43}$$

where p is the number of variables and n is the number of samples [55]. If the data is in fact multivariate normally distributed, then the $\mathbb{E}[b_{1,p}] = 0$ and $\mathbb{E}[b_{2,p}] = p(+2)$ [57]. From [54], a function of multivariate skewness is asymptotically distributed as a χ^2 random variable with $\frac{p(p+1)(p+2)}{6}$ degrees of freedom and the multivariate kurtosis is asymptotically distributed as a standard normal random variable with a mean of $p(p+2)$ and a variance of $\frac{8p(p+2)}{n}$.

4.4.1.2 Royston's H Test.

Common tests for univariate normality are the Shapiro-Wilk test [79] and the Kolmogorov-Smirnov test [56]. Royston extended the Shapiro-Wilk W test in such a way as to provide good results [58] when applied to the multivariate case [76].

Roystons H test begins by using Shapiro-Wilk univariate normality test on each of the n variates. The second part of the test combines the dependent tests into one test statistic for multivariate normality which is approximately a χ^2 random variable.

Letting W_i be the Shapiro-Wilk statistic for the ith variable in a p-variate distribution, and then define

$$R_i = \left\{ \Phi^{-1} \left[\frac{1}{2} \Phi \left\{ -((1-W_i)^\lambda - \mu)/\sigma \right\} \right] \right\}^2, \tag{44}$$

where λ, μ, and σ are calculated from polynomial approximations that Royston gives in [75]. The variables Φ^{-1} and Φ are the inverse and standard normal cumulative distribution function, respectively.

If the data are from a multivariate normal distribution, then

$$H = \xi \sum_{i=1}^{p} R_i / p \tag{45}$$

approximates a χ^2 distribution, where the equivalent degrees of freedom are

$$\hat{\xi} = \frac{p}{[1 + (p-1)\hat{c}]}, \tag{46}$$

$\hat{c}$ is an estimate of the average correlation among the R_j's [76]. This χ^2 distribution is used to obtain the p-values for the multivariate normal test [76].

4.4.1.3 Henze-Zirkler Test.

Another powerful multivariate normal test is the the Henze-Zirkler test. It is based on a nonnegative functional distance that measures the distance between two distribution functions: the characteristic function of the multivariate normality and the empirical characteristic function.

The Henze-Zirkler test statistic is defined as follows. Let $\mathbf{X}_1$, $\mathbf{X}_2$,...,$\mathbf{X}_n$ be a set of n independent data samples (i.e., the $\mathbf{X}_1$, $\mathbf{X}_2$,...,$\mathbf{X}_n$ are obtained from n independent records) each of dimension d (i.e., $\mathbf{X}_i = \{x_{i_1}, x_{i_2}, ..., x_{i_d}\}$). A closed form of the test statistic is given as:

$$\begin{aligned} T_{n,b} = & \frac{1}{n} \sum_{k=1}^{n} \sum_{j=1}^{n} \left[exp\left(-\frac{\beta^2}{2} ||Y_j - Y_k||^2 \right) \right] \\ & -2(1+\beta^2)^{-d/2} \sum_{j=1}^{n} \left[exp\left(-\frac{\beta^2}{2(1+\beta^2)} ||Y_j||^2 \right) \right] + n(1+2\beta^2)^{-d/2} \end{aligned} \tag{47}$$

where

$$\begin{aligned} \beta &= \frac{1}{\sqrt{2}} \left(\frac{2d+1}{4} \right)^{\frac{1}{d+4}} n^{\frac{1}{d+4}}, \\ ||Y_j - Y_k||^2 &= \left(\mathbf{X}_j - \mathbf{X}_k \right)^T \mathbf{S}^{-1} \left(\mathbf{X}_j - \mathbf{X}_k \right), \\ ||Y_j||^2 &= \left(\mathbf{X}_j - \bar{\mathbf{X}}_n \right)^T \mathbf{S}^{-1} \left(\mathbf{X}_j - \bar{\mathbf{X}}_n \right), \end{aligned}$$

$T_{n,\beta}$ is the test statistic, $\bar{\mathbf{X}}_n$ is the sample mean vector of the n realizations $\mathbf{X}_1$, $\mathbf{X}_2$,...,$\mathbf{X}_n$, and $\mathbf{S}$ is the sample covariance matrix defined as

$$\mathbf{S} = \frac{1}{n}\sum_{j=1}^{n}\left(\mathbf{X}_j - \bar{\mathbf{X}}_n\right)\left(\mathbf{X}_j - \bar{\mathbf{X}}_n\right)^T \tag{48}$$

The Henze-Zirkler statistic is approximately distributed as a log-normal. The log-normal distribution is used to compute the null hypothesis probability.

According to Henze-Wagner [37], their test has the following desirable properties: (a) affine invariant; (b) consistency against each fixed non-normal alternative distribution; (c) asymptotic power against contiguous alternatives of order $n^{-1/2}$; and (d) feasibility for any dimension and any sample size.

4.4.2 Discussion on Normality Testing Results.

A note about all of the multivariate tests that use the p-value. If the p-value of any of the tests is small, then multivariate normality will be rejected. However, the univariate Shapiro-Wilk W test is a very powerful test and is capable of detecting small departures from univariate normality with relatively small sample sizes, thus causing the test to reject univariate normality (and therefore multivariate normality) unnecessarily, because the results upon which the multivariate test is being done may be from a robust system insensitive to small departures from normality [58].

Based on the nature of the algorithm's design for the multivariate tests, each test returns a true or false value based upon whether the null hypothesis of multivariate normality is a reasonable assumption based on a specified significance level, often 5%. When these tests are conducted on the results of a Monte-Carlo simulation, there is high confidence that the distributions for the head, hips, shoulders, ankles and elbows are generally multivariate normal. However, the tests indicate that, in general, the knee and wrist markers are not multivariate normal with the distributions for those markers often indicating bimodality. Fig. 45 provides the histograms for a random subject. For this subject, the distributions of both knees, the left hip, and the left wrist are clearly bimodal.

4.4.3 Mahalanobis Distance.

A useful measure often employed to evaluate multivariate data is the Mahalanobis distance. Its power derives from the fact that it takes into account the covariance of the distribution for the variables in question, eliminating the problems that may arise from using the Euclidean distance, namely scale and correlation. A multivariate normal distribution's isocontours might not be spherical nor might the principal axes of the formed hyperellipsoid coincide with the axes of the Cartesian coordinate system, thus assuming the Euclidean distance from the mean holds the same importance throughout the data set may be erroneous. Since the Mahalanobis distance relies on the distribution of the data, it becomes an extremely useful determinant in detecting outlying points and judging classification [34]. The standard equation for the squared Mahalanobis distance between $\mathbf{x}$ and μ [18] is defined as:

$$r^2 = (\mathbf{x} - \mu)^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \mu). \tag{49}$$

Calculating the Mahalanobis distance for each set of estimated marker locations to the means of the estimated marker set from the Monte Carlo simulation and looking at the histograms of the distances provides the results shown in Fig. 56. The calculation is only valid for distributions that are multivariate normal; therefore, bimodal marker distributions are not presented in Fig. 56. Each plot also has vertical lines representing the σ, 2σ, and 3σ distances.

Examining the histograms of the Mahalanobis distances indicates that there are outlying points for most of the sets of markers. However, in the case of the distribution for the left knee markers, which was shown by the Mardia, Royston, and Henze-Zirkler tests as not being multivariate normal, the outlying point seems to be just that, an outlier. Whereas, the majority of the calculated left knee markers fall within the bounds of a multivariate normal distribution indicating that assuming a multivariate normal distribution is not a bad assumption.

We can also look at the Mahalahobis distance for each set of estimated marker locations to the set of truth markers which provides an idea of the performance of the algorithm at

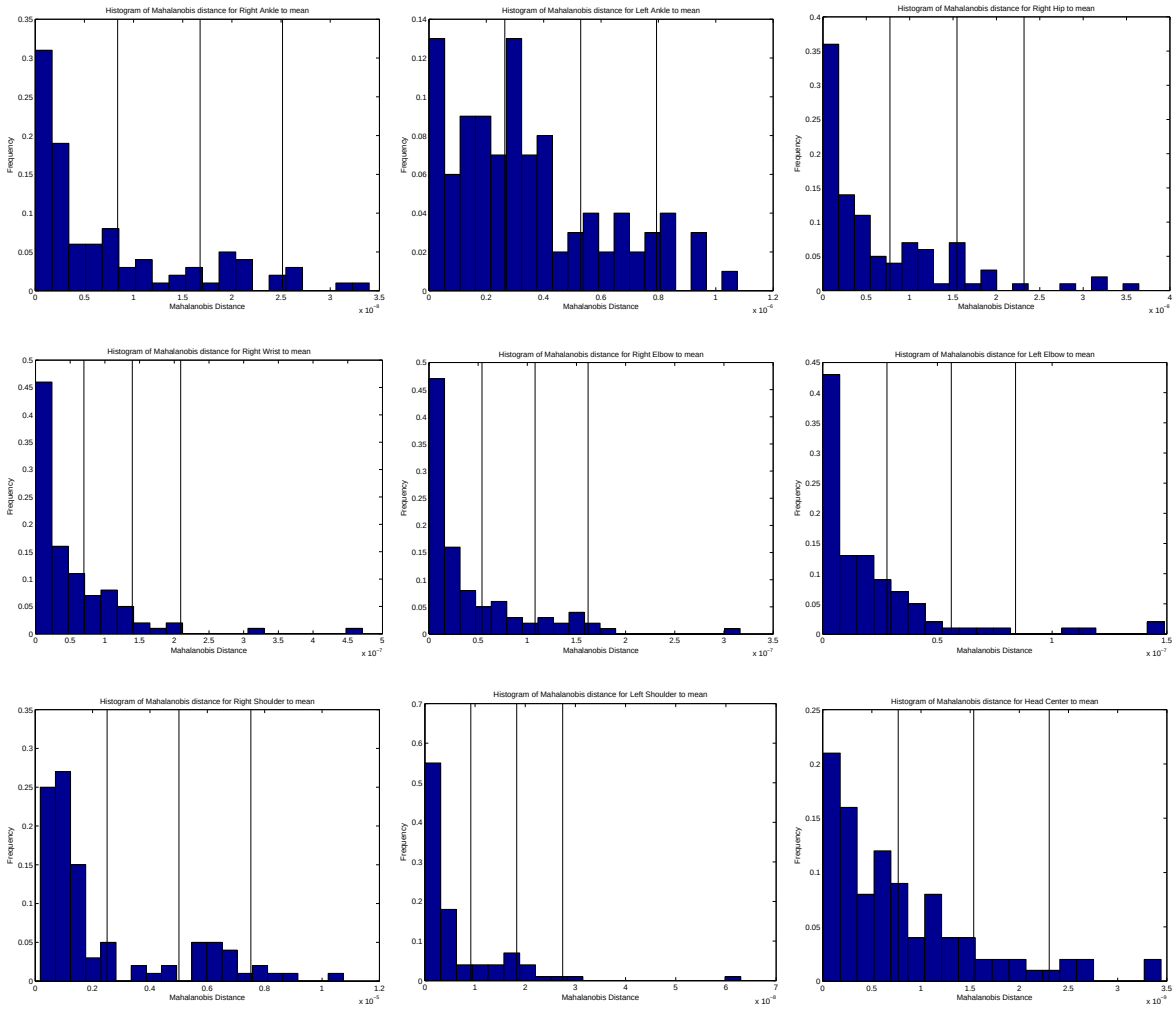


Figure 56. Histograms of Mahalanobis distances to mean estimated marker locations for 100 Monte Carlo simulations with the σ, 2σ, and 3σ bounds indicated as vertical lines on the plot.

estimating each marker. Calculating the Mahalanobis distances and plotting the histograms of the distances provides the results shown in Fig. 57. Each plot also has vertical lines representing the σ, 2σ, and 3σ distances.

Examining the Mahalanobis distances shown in Fig. 57 provides information about the systematic biases produced by the marker estimation algorithm. Additionally, insight into how well a multivariate normal distribution fits the estimated markers is gained by examining the distribution of the Mahalanobis distances. As with univariate normal distributions, we expect a majority of the Mahalanobis distances to fall within 1σ of the truth marker. When this is not the case, a systematic bias is observed.

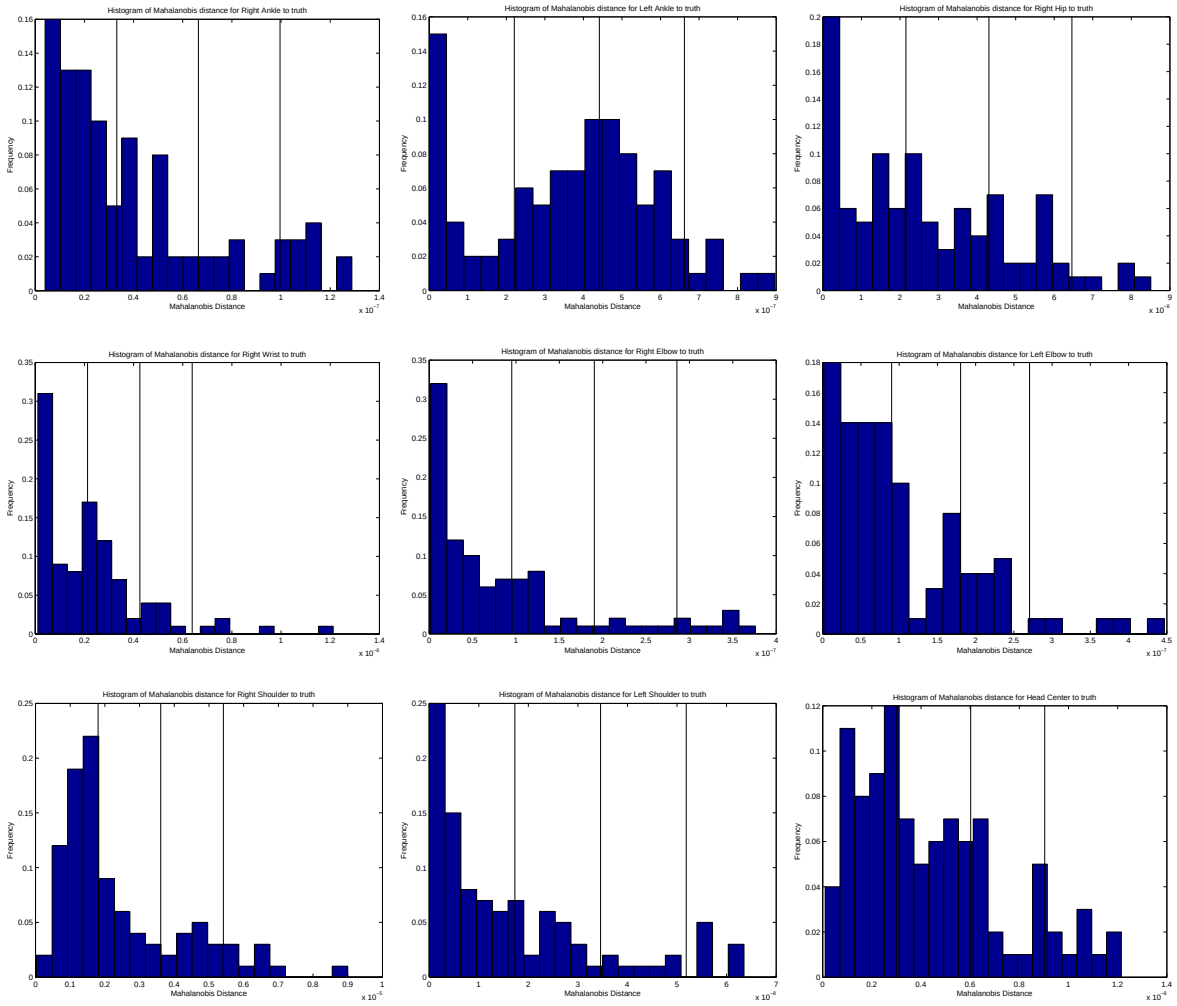


Figure 57. Histograms of Mahalanobis distances to truth markers for 100 Monte Carlo simulations with the σ, 2σ, and 3σ bounds indicated as vertical lines on the plot.

Another measure worth examining is the mean Euclidean distance. The mean Euclidean distance provides information about the accuracy of the estimates as well as insight into the systematic biases. Table 9 provides the mean Euclidean distances, in the re-scaled space, of the estimated markers from the Monte Carlo simulations to the truth marker. Additionally, Table 9 provides the span of the distances, the median distance, and the consistency of the distances as indicated by the variance. Examining these results, we see that the estimated hip markers have the largest average distance from the truth markers. The largest variances correspond to the hip and knee centers, which is an expected result based on the distribution of estimated markers shown in Fig. 45.

Table 9. Statistics for the Euclidean distance of estimated markers from Monte Carlo simulation to truth markers.

Marker Set	Mean (mm)	Minimum (mm)	Maximum (mm)	Median (mm)	Variance (mm^2)
Right Ankle	37.40	15.75	68.82	33.52	202.91
Left Ankle	39.51	9.70	73.05	36.20	197.66
Right Knee	38.45	0.64	66.92	44.35	290.40
Left Knee	38.34	1.83	65.82	43.53	281.34
Right Wrist	20.59	10.74	37.68	18.78	44.38
Right Hip	75.09	42.67	134.31	69.47	443.65
Left Hip	63.72	10.11	128.80	57.56	544.97
Left Wrist	16.65	2.91	37.73	14.21	91.28
Right Elbow	28.46	12.73	53.72	26.42	70.64
Left Elbow	29.17	12.79	66.29	26.71	136.39
Right Shoulder	32.99	16.11	76.10	32.08	120.76
Left Shoulder	32.27	15.35	51.38	31.36	54.82
Head Center	21.73	3.49	39.60	21.37	65.33
All Distances	36.49	0.64	134.31	31.66	444.99

4.5 Classification Results

4.5.1 Case I - Truth Data versus Truth Data.

Here, the truth data is used as the training data and the testing data for the classification algorithm. The results are provided in Table 10 for the classification accuracy using resubstitution and 10-fold cross validation. As discussed in Sections 4.2.2 and 4.2.3, classification is done on the entire database and also on the three countries separately. As expected, the classification results for the reduced feature set are worse than the results obtained in [27], where the 16-measurement feature set is used. Furthermore, 10-fold cross validation produces a lower classification accuracy than resubstitution. The reduction in classification accuracy for K-fold cross validation is consistent with the expected results from classification theory [34]. An additional observation is that the reported results in [27] show a diminished classification accuracy for the Netherlands portion of the CAESAR database which holds true when using the reduced feature set. In [27], the cause for this reduction in accuracy for the Netherlands database is surmised to be due to differences in the scanner's calibration, resolution, and processing.

Table 10. Classification accuracy for the truth data using linear discriminant analysis (LDA).

Scan Location	Resubstitution	10-fold Cross-Validation
North America	89.5%	89.3%
Italy	94.8%	86.0%
Netherlands	87.3%	87.1%
Entire Database	88.7%	88.57%

4.5.2 Case II - Estimated Markers versus Truth Data.

Using the truth data as the training set for the LDA classification and the estimated markers as the test set produces a classification accuracy of 60.8%. The test set consists of an approximately equal number of male and female subjects from each separate section of the database. Some further insight into the problem is seen when looking at the classification accuracy of each nation's database separately. As shown in Table 11, the result of a lower classification accuracy for the Netherlands portion of the CAESAR database is consistent with the baseline results established in Case I.

Table 11. Classification accuracy for the calculated marker locations versus the truth data.

Scan Location	Classification Accuracy
North America	62.9%
Italy	69.3%
Netherlands	55.2%
Entire Database	60.8%

4.5.3 Case III - Monte-Carlo Means versus Monte-Carlo Means.

In this case, the test and training data come from the same data set, the mean marker locations produced from the Monte Carlo simulation of each subject. Similar to the classification results shown in Table 10, the classification process calculates the resubstitution accuracy and estimates the true generalization error through K-fold cross-validation. The results are provided in Table 12.

Table 12. Classification accuracy for the means of the Monte Carlo simulation as both training and testing data.

Scan Location	Resubstitution	10-fold Cross Validation
North America	83.9%	83.1%
Italy	85.2%	83.2%
Netherlands	82.3%	81.3%
Entire Database	77.1%	77.2%

4.5.4 Case IV - Estimated Markers versus Monte-Carlo Means.

The test data consists of a set of estimated marker locations from a single run of the marker estimation algorithm, and the training data is comprised of the mean marker locations produced by the Monte Carlo simulation using ten runs of the marker estimation algorithm for each subject. The test set consists of an approximately equal number of male and female subjects drawn from each subsection of the database. The results are presented in Table 13. As opposed to the results in Table 11, the reduction in classification accuracy when using a separate test set versus resubstitution and K-fold cross validation is minimal.

Table 13. Classification accuracy for test data consisting of estimated marker locations classified against the training data consisting of the means of the Monte-Carlo simulation runs.

Scan Location	Classification Accuracy
North America	79.6%
Italy	79.7%
Netherlands	74.2
Entire Database	74.3%

4.6 Analysis of Results and Summary

The performance of the LDA classification algorithm itself is confirmed as a suitable method for this problem by the duplication of the results from [27]. The low classification accuracy of the calculated markers versus the truth data in Case II is indicative of systematic errors the methodology injects into the calculated markers. This supposition is supported by improved classification accuracy obtained when classifying calculated markers against the average marker locations of the Monte Carlo simulations in Case IV. The consistently reduced classification accuracies for the Netherlands database are an interesting observation,

but there are no further insights to be added to those mentioned in [27]. Overall, the 74.3% classification accuracy presented in Table 13 indicates that the novel approach developed in this thesis for anthropometric marker estimation performs well in segmenting the 3-D LADAR point cloud data.

In this chapter, the simulation process, parameter selection, and segmentation results are presented. The gender classification results are provided, indicating that the algorithm is a promising method for gender classification from point cloud data. In the next chapter, the results are summarized and areas for future work are described.

5. Conclusions and Future Work

This chapter begins with a discussion of methods used to achieve the research's goal, followed by a summary of the results, and conclusions. The final section covers the potential for future work to improve the algorithm inspected here.

5.1 Summary of Methods and Conclusions

The goal of this thesis is to describe an innovative segmentation algorithm that extracts anthropometric markers from a 3-D LADAR point cloud. The algorithm is built on the dimensionality reduction tool, principal component analysis (PCA), as the initial step. Clustering is performed using a Self-Organizing Map (SOM) to explore the distribution of the data. Anthropometric knowledge in conjunction with a skeletal framework obtained by the use of an alpha hull is incorporated to refine the location of the markers obtained from the SOM map. The validity of the marker locations is then tested using linear discriminant analysis (LDA) as the classification tool, following [27].

The quality of the estimation algorithm is measured by performing gender classification on the resulting marker set. The performance of the algorithm was compared to the true marker location with limited success, with an accuracy of 61%. However, this is not entirely indicative of how well the algorithm performs since truth data is not available in real-world scenarios. When a set of average marker locations, obtained from Monte Carlo simulations, is used as the training set, the classification accuracy increases to approximately 74%, a significant accuracy for the first algorithm to extract anthropometric markers from 3-D point cloud data. This result indicates that the algorithm may show promising results when applied to a more realistic data set, lending credence to the notion that the avenue of marker extraction developed in this thesis will serve as a useful addition to the research group's dismount detection system when the sensor fusion effort grows to include a LADAR system.

5.2 Future Work

This research effort indicates the feasibility of successfully extracting anthropometric markers from a 3-D LADAR point cloud. Despite the evidence of this success, there are still a variety of avenues in which the performance of the marker extraction algorithm may be improved.

Certainly, a major area for improvement is in developing a richer data set for training the system. A data set that contains subjects from more diverse backgrounds would help in determining if other demographic factors can be determined from the markers extracted from a 3-D point cloud. Additionally, a data set comprised of different poses would prove beneficial in examining the performance of the algorithm and discovering where modifications need be made.

Directly related to the decisions made in implementing the algorithm, further research should be done to examine the effects of hip marker placement to the overall classification accuracy. Specific decisions were made during the algorithm development process in linking the location of the hip markers linearly to the ankle-knee marker line. This decision may not accurately predict marker locations for subjects who are knock-kneed or bow-legged. A comparison study should be between allowing the hip centers to remain tied to the results of the SOM with minimal anthropometric refinement and using the current slope-adjusted method.

Similarly, a comparison study could be conducted using a modified marker estimation algorithm. Near the end of the research effort, work was begun on a method that removed the SOM step from the algorithm, replacing it with an approximate marker estimate along the x-axis based on the golden segment principles and snapped to the alpha hull derived skeleton. The resulting algorithm is significantly faster. However, its performance was not evaluated using the developed classification algorithm. Conducting this performance evaluation may provide some insight into the importance of the SOM in the developed approach.

An additional area for further investigation is in comparing the performance of the

marker extraction algorithm using other classification methods. For example, one may consider a backpropogation neural network, support vector machines, or other statistical methods.

An intriguing and vital area for future work is in exploring the performance of the algorithm when the data comes from a Flash LADAR system and when confusers are added to the scene. This may not be practical until the data fusion step occurs since the process may need to trigger off of a valid skin detection from the previous work into dismount detection using the multispectral camera system [8, 14, 15, 45, 65, 67].

5.3 Contributions

The research addressed in this thesis proposes a new approach to the problem of segmenting 3-D LADAR point cloud data. The classification results indicate the marker estimation algorithm successfully locates anthropometric markers, and the resulting gender information is a valid addition to dismount characterization. The lasting results from this thesis are that extracting anthropometric markers from point cloud data is possible with a great deal of success and the results are significant enough to warrant incorporation into the research group's dismount detection and characterization system.

Appendix A. Mahalanobis Calculations for Another Subject

This appendix shows the multivariate analysis results for another subject in Fig. 58, Fig. 59, Fig. 60, Fig. 61, and Fig. 62.

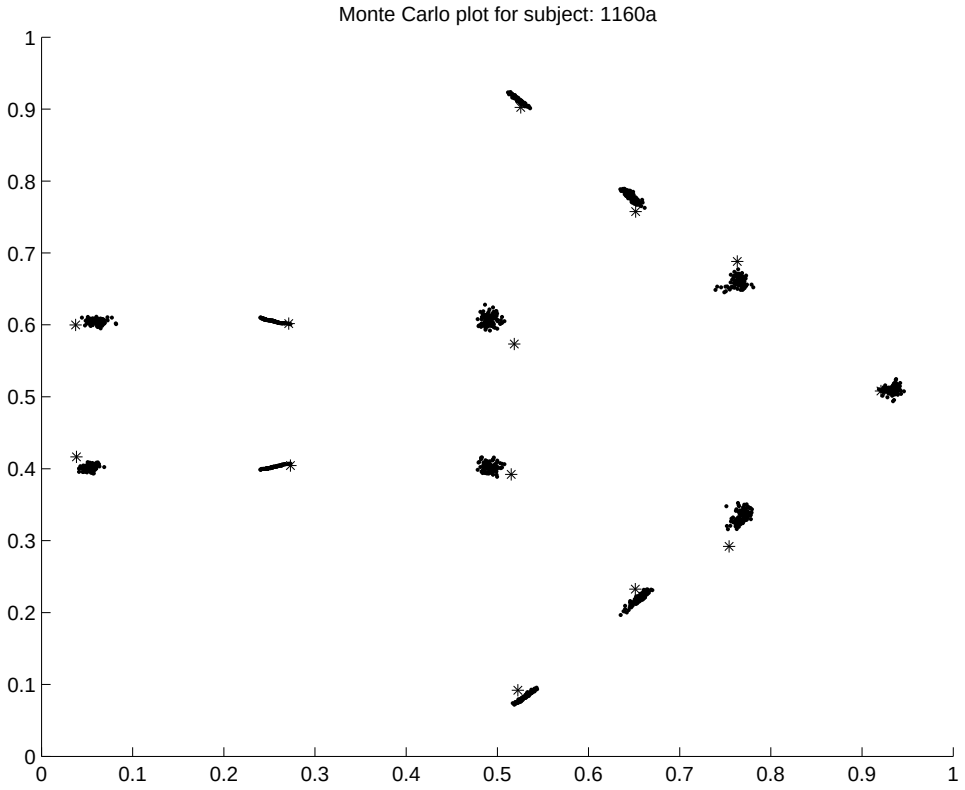


Figure 58. An example of the distribution of calculated markers from the Monte Carlo simulation for a random subject.

Since none of the histograms clearly indicate bimodality, all of the Mahalanobis distance computations for the subject are presented in Figs. 61 and 62.

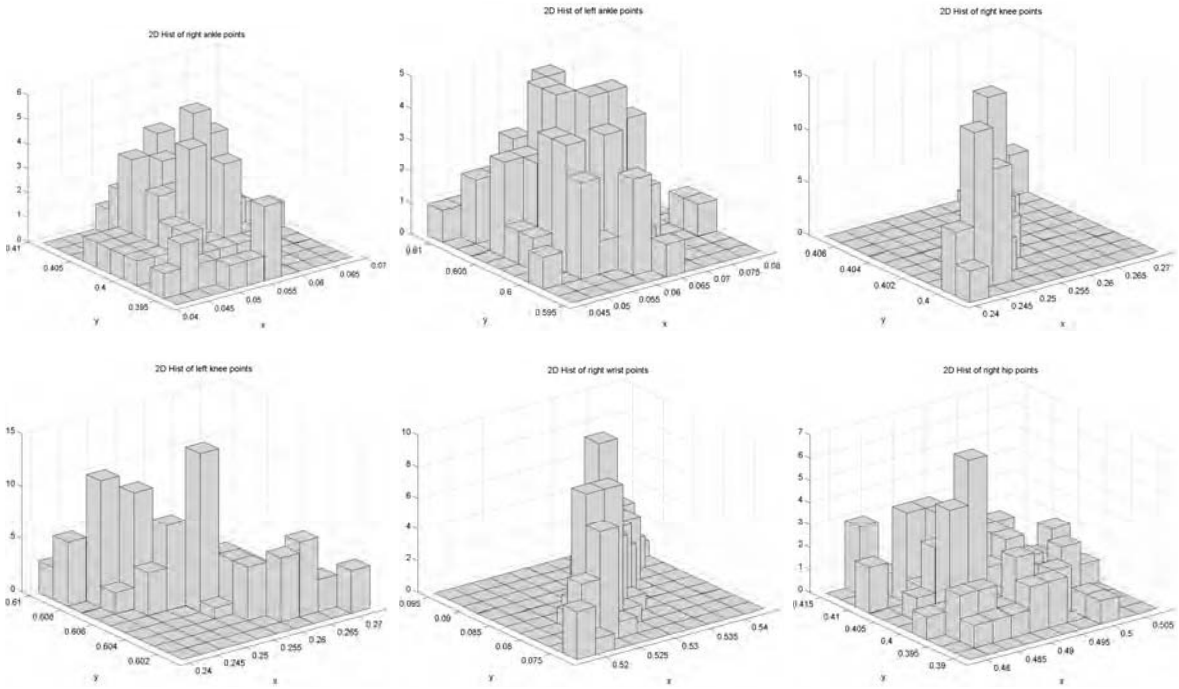


Figure 59. Part 1 - Histograms of the 100 run Monte Carlo simulations.

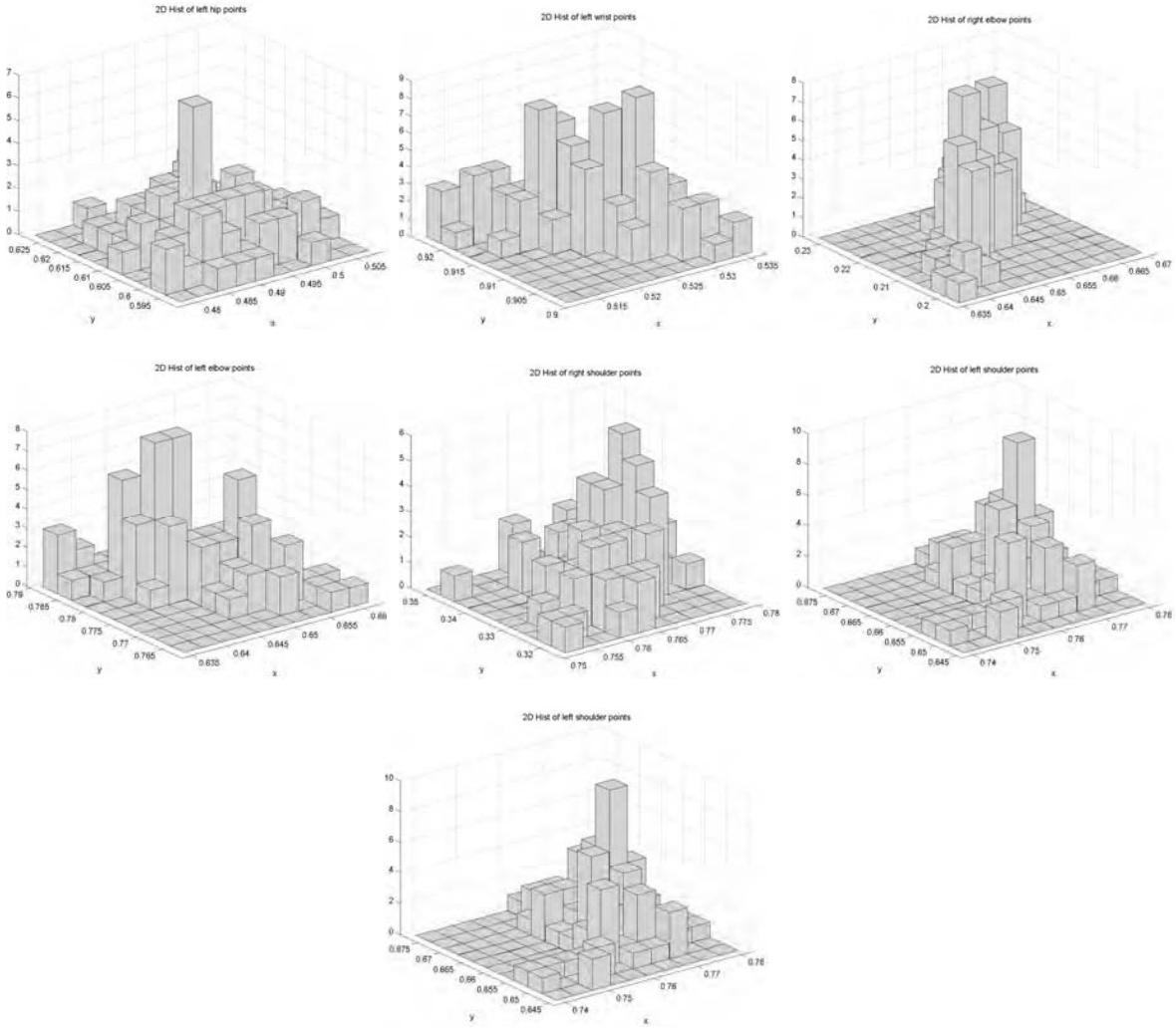


Figure 60. Part 2 - Histograms of the 100 run Monte Carlo simulations.

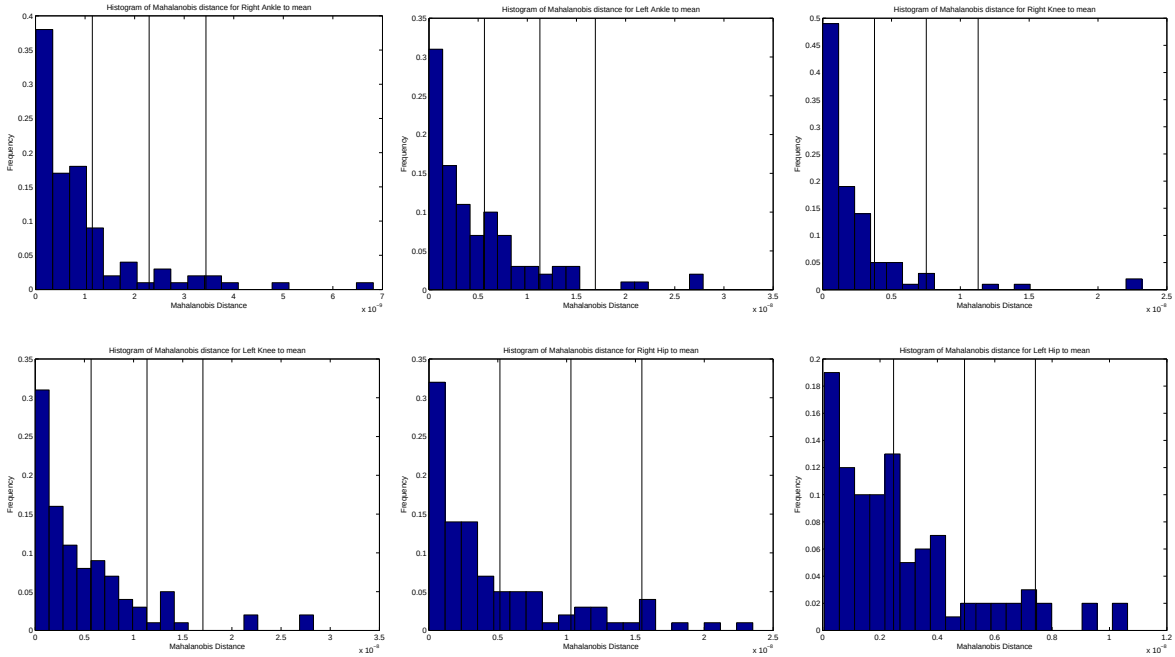


Figure 61. Part 1 - Histograms of Mahalanobis distances to mean estimated marker locations for 100 Monte Carlo simulations with the σ, 2σ, and 3σ bounds indicated as vertical lines on the plot.

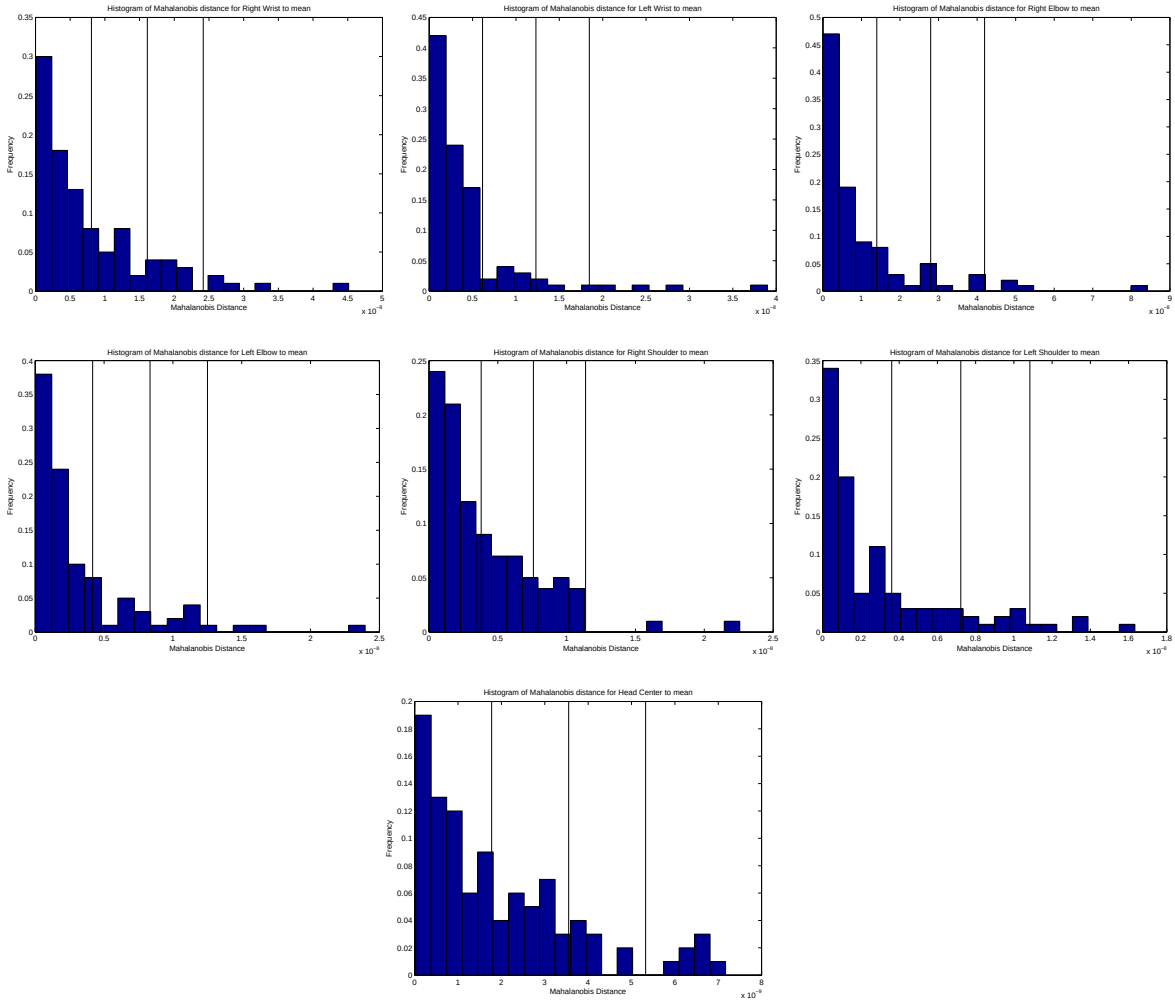


Figure 62. Part 2 - Histograms of Mahalanobis distances to mean estimated marker locations for 100 Monte Carlo simulations with the σ, 2σ, and 3σ bounds indicated as vertical lines on the plot.

Appendix B. Software Acknowledgments

In the course of performing some of the multivariate tests in this research effort several MATLAB scripts that implemented the MVN tests were used. Among these are the mardiatest [33], HZmvntest [83], and Roystest [84]. The ply file reader developed by Pascal Getreur [30] was used to easily read in the 3-D LADAR body scans.

Bibliography

[1] Attali, D., H. Edelsbrunner, and Y. Mileyko. "Weak witnesses for Delaunay triangulations of submanifolds". *Proceedings of the 2007 Association for Computing Machinery Symposium on Solid and Physical Modeling*, Solid and Physical Modeling 2007, 143–150. ACM, New York, NY, USA, 2007. ISBN 978-1-59593-666-0. URL `http://doi.acm.org/10.1145/1236246.1236267`.

[2] Attene, M., B. Falcidieno, and M. Spagnuolo. "Hierarchical mesh segmentation based on fitting primitives". *Vis. Comput.*, 22:181–193, March 2006. ISSN 0178-2789. URL `http://portal.acm.org/citation.cfm?id=1127195.1127199`.

[3] Belkin, M. and P. Niyogi. "Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering". *Advances in Neural Information Processing Systems 14*, 585–591. MIT Press, 2001.

[4] Biasotti, S. "Topological Techniques for Shape Understanding". *In Central European Seminar on Computer Graphics, CESCG*. 2001.

[5] Bishop, C.M. *Neural networks for pattern recognition*. Oxford University Press, USA, 1995.

[6] Blaak, E. "Gender differences in fat metabolism." *Curr Opin Clin Nutr Metab Care*, 4(6):499–502, Nov 2001.

[7] Blum, H. "A Transformation for Extracting New Descriptors of Shape". Weiant Wathen-Dunn (editor), *Models for the Perception of Speech and Visual Form*, 362–380. MIT Press, Cambridge, 1967.

[8] Brooks, A. *Improved multispectral skin detection and its application to search space reduction for dismount detection based on histograms of oriented gradients*. Master's thesis, Air Force Institute of Technology, 2010.

[9] Brostow, Gabriel J., Irfan Essa, Drew Steedly, and Vivek Kwatra. "Novel skeletal representation for articulated creatures". *In Proc. European Conf. on Computer Vision*, 66–78. 2004.

[10] Camomilla, V., Cereatti. A., G. Vannozzi, and A. Cappozzo. "An optimized protocol for hip joint centre determination using the functional method." *J Biomech*, 39(6):1096–1106, 2006. URL `http://dx.doi.org/10.1016/j.jbiomech.2005.02.008`.

[11] Chehata, N. and F. Bretar. "Terrain modeling from lidar data: Hierarchical K-means filtering and Markovian regularization". *Proc. 15th IEEE Int. Conf. Image Processing ICIP 2008*, 1900–1903. 2008.

[12] Chehata, N., N. David, and F. Bretar. "LIDAR Data Classification using Hierarchical K-means clustering". *Management*, 37(2000):207–212, 2008. URL `http://www.isprs.org/congresses/beijing2008/proceedings/3b_pdf/65.pdf`.

[13] Chun, C., O. C. Jenkins, and M. J. Mataric. "Markerless kinematic model and motion capture from volume sequences". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2. 2003.

[14] Clark, J. *Stochastic Feature Selection with Distributed Spacing and its Application to Textile*. Ph.d. prospectus, Air Force Institute of Technology, 2010.

[15] Climer, J. *Overcoming Pose Limitations of a Skin-Cued Histograms of Oriented Gradients Dismount Detector through Contextual Use of Skin Islands and Multiple Support Vector Machines*. Master's thesis, Air Force Institute of Technology, 2011.

[16] Cuzzolin, F., D. Mateus, D. Knossow, E. Boyer, and R.P. Horaud. "Coherent Laplacian 3-D Protrusion Segmentation". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2008. URL `http://perception.inrialpes.fr/Publications/2008/CMKBH08`.

[17] Da Vinci, L. *The Notebooks of Leonardo Da Vinci Complete*. Public Domain, 2004.

[18] Duda, R., P. Hart, and D. Stork. *Pattern Classification*. Wiley, New York, 2. edition, 2001.

[19] Edelsbrunner, H. and J. Harer. *Computational Topology, An Introduction*. American Mathematical Society, January 2010. ISBN 0-8218-4925-5.

[20] Edelsbrunner, J., D. Kirkpatrick, and R. Seidel. "On the shape of a set of points in the plane". *IEEE Transactions on Information Theory*, 29:551–558, 1983.

[21] Ester, M., H. Kriegel, J. Sander, and X. Xu. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". *KDD*, 226–231. 1996.

[22] Fabbri, R., L. F. Estrozi, and L. F. Costa. "On Voronoi Diagrams and Medial Axes". *Journal of Mathematical Imaging and Vision*, 17:27–40, 2002.

[23] Fausett, L. *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1994. ISBN 0-13-334186-0.

[24] Fodor, I. "A Survey of Dimension Reduction Techniques", 2002. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.8.5098`.

[25] Frank, A. and A. Asuncion. "UCI Machine Learning Repository", 2010. URL `http://archive.ics.uci.edu/ml`.

[26] Fullenkamp, A. M. "Anthropometric discussion and e-mails", Jan 2011.

[27] Fullenkamp, A. M., K. Robinette, and H. A. M. Daanen. "Gender differences in NATO anthropometry and the implication for protective equipment". *in NATO Research and Technology Organisation (RTO) Human Factors and Medicine Panel (HFM) Symposium*. 2008.

[28] Funatomi, T., I. Moro, S. Mizuta, and M. Minoh. "Surface Reconstruction from Point Cloud of Human Body by Clustering". *Systems and Computers in Japan*, 37:44–56, October 2006. URL `http://portal.acm.org/citation.cfm?id=1150833.1150839`.

[29] Gavrila, D. M. and L. S. Davis. "3-D model-based tracking of humans in action: a multi-view approach". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '96*, 73–80. 1996.

[30] Getreuer, P. "Ply File Manipulating Functions". WWW Document. MATLAB Files, 2004.

[31] de Goes, Fernando, Siome Goldenstein, and Luiz Velho. "A hierarchical segmentation of articulated bodies". *Proceedings of the Symposium on Geometry Processing*, Symposium on Geometry Processing 20'08, 1349–1356. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 2008. URL `http://portal.acm.org/citation.cfm?id=1731309.1731315`.

[32] Grabusts, P. "A Study of Clustering Algorithm Application in RBF Neural Networks", 2001.

[33] Graham, D. "Mardiatest: A Matlab File". WWW Document, 2006. Loughborough University.

[34] Hastie, T., R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

[35] Haykin, S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Upper Saddle River, NJ, 1999. 2nd edition.

[36] Hegde, C., M.B. Wakin, and R.G. Baraniuk. "Random Projections for Manifold Learning". John C. Platt, Daphne Koller, Yoram Singer, and Sam T. Roweis (editors), *Neural Information Processing Systems (NIPS)*. MIT Press, 2007.

[37] Henze, N. and T. Wagner. "A new approach to the BHEP tests for multivariate normality". *Journal of Multivariate Analysis*, 62:1–23, July 1997. ISSN 0047-259X. URL `http://portal.acm.org/citation.cfm?id=261794.261796`.

[38] Hoover, A., G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher. "An experimental comparison of range image segmentation algorithms". *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 18(7):673–689, 1996.

[39] Hsu, A. L. and S. K. Halgarmuge. "Enhanced Topology Preservation of Dynamic Self-Organising Maps for Data Visualisation". *Proceedings of the Joint 9th International Fussy Systems Association (IFSA) World Congress and 20th Norht American Fuzzy Information Processing Society (NAFIPS) International Conference*, volume 3, 1786–1791. 2001.

[40] Jaggers, T.J. "Presentation to the House Armed Forces Committee Subcommittee on Terrorism, Unconventional Threats and Capabilities", 2009.

[41] Jain, A. K. *Fundamentals of Digital Image Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989. ISBN 0-13-336165-9.

[42] Kang, M. "Hip Joint Center Location by Fitting Conchoid Shape to the Acetabular Rim Region of MR Images." *Conference Proceedings of the IEEE Engineering in Medicine and Biology Society*, 6:4477–4480, 2004. URL `http://dx.doi.org/10.1109/IEMBS.2004.1404244`.

[43] Kiviluoto, K. "Topology preservation in self-organizing maps". *International Symposium on Neural Networks*. 1996.

[44] Klasing, K., D. Wollherr, and M. Buss. "A clustering method for efficient segmentation of 3d laser data". *in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2008.

[45] Koch, B. *A Multispectral Bidirectional Reflectance Distribution Function Study of Human Skin for Improved Dismount Detection*. Master's thesis, Air Force Institute of Technology, 2011.

[46] Kohonen, T. *Self-Organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 3rd edition, 2001. ISBN 3540679219.

[47] Kolluri, R., J. R. Shewchuk, and J. F. O'Brien. "Spectral Surface Reconstruction from Noisy Point Clouds". *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry Processing*, Symposium on Geometry Processing 2004, 11–21. ACM, New York, NY, USA, 2004. ISBN 3-905673-13-4. URL `http://doi.acm.org/10.1145/1057432.1057434`.

[48] Koschan, A. F. "Perception-based 3-D Triangle Mesh Segmentation Using Fast Marching Watersheds". *in Proceedings of the International Conference on Computer Vision and Pattern Recognition, II*, 27–32. 2003.

[49] Litinskii, L.B. and D.E. Romanov. "Neural Network Clustering Based on Distances Between Objects". *International Conference on Artificial Neural Networks*, 437–443. 2006.

[50] Liu, R. and H. Zhang. "Segmentation of 3D meshes through spectral clustering". *Proceedings of the 12th Pacific Conference on Computer Graphics and Applications*, Pacific Graphics 2004, 298–305. 2004.

[51] Liu, X., S. Zhong, and X. Bai. "A Modified SOFM Segmentation Method in Reverse Engineering". *Proceedings of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing - Volume 02*, SNPD '07, 570–573. IEEE Computer Society, Washington, DC, USA, 2007. ISBN 0-7695-2909-7. URL `http://dx.doi.org/10.1109/SNPD.2007.46`.

[52] Mangan, A.P. and R.T. Whitaker. "Partitioning 3D surface meshes using watershed segmentation". *IEEE Transactions on Visualization and Computer Graphics*, 5(4):308–321, 1999. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=817348`.

[53] Marchette, D.J. *Random Graphs for Statistical Pattern Recognition*. Wiley-Interscience, 2004. ISBN 0471221767.

[54] Mardia, K. V. "Measures of multivariate skewness and kurtosis with applications". *Biometrika*, 57(3):519–530, 1970. URL http://biomet.oxfordjournals.org/cgi/doi/10.1093/biomet/57.3.519.

[55] Mardia, K. V., J.T. Kent, and J.M. Bibby. *Mutivariate Analysis*. Academic Press, 1979.

[56] Massey, F.J. "The Kolmogorov-Smirnov Test for Goodness of Fit". *Journal of the American Statistical Association*, 46(253):68–78, 1951. URL http://www.jstor.org/stable/2280095.

[57] Mecklin, C. and D. Mundfrom. "Comparing the Power of Classical and Newer Tests of Multivariate Normality". *81st Annual Meeting of the American Educational Research Association*. 2000.

[58] Mecklin, C.J. *A Comparison of the Power of Classical and Newer Tests of Multivariate Normality*. Ph.D. thesis, University of Northern Colorado, 2000.

[59] Mecklin, C.J. and D.J. Mundfrom. "A Monte Carlo comparison of the Type I and Type II error rates of tests of multivariate normality". *Journal of Statistical Computation and Simulation*, 75:93–107, 2005.

[60] Mederos, B., N. Amenta, L. Velho, and L. de Figueiredo. "Surface reconstruction from noisy point clouds". *Proceedings of the Third Eurographics Symposium on Geometry Processing*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 2005. ISBN 3-905673-24-X. URL http://portal.acm.org/citation.cfm?id=1281920.1281929.

[61] Memarsadeghi, N., D.M. Mount, N.S. Netanyahu, and J.L. Moigne. "A Fast Implementation of the Isodata Clustering Algorithm". *International Journal of Computational Geometry and Applications*, 17(1):71–103, 2007.

[62] Mortara, M. and G. Patané. "Affine-Invariant Skeleton of 3D Shapes". *Proceedings of the Shape Modeling International 2002 (SMI'02)*, 245–. IEEE Computer Society, Washington, DC, USA, 2002. ISBN 0-7695-1546-0. URL http://portal.acm.org/citation.cfm?id=882487.884159.

[63] Motoc, A., M. Motoc, S. Bolintineanu, C. Muşuroi, and M. Munteanu. "The construction of human body - from model to reality." *Romanian Journal of Morphology and Embryology*, 46(1):63–6, 2005.

[64] Ng, R.T. and J. Han. "CLARANS: A Method for Clustering Objects for Spatial Data Mining". *IEEE Transactions on Knowledge and Data Engineering*, 14:1003–1016, September 2002. ISSN 1041-4347. URL http://dx.doi.org/10.1109/TKDE.2002.1033770.

[65] Nunez, A. S. *A physical model of human skin and its application for search and rescue*. Ph.D. thesis, Air Force Institute of Technology, 2009.

[66] Park, S., S. Lee, and H. Choi. "Multiscale surface representation and rendering for point clouds". *Proceedings of the International Conference on Image Processing (ICIP) 2004*, volume 3, 1939–1942. 2004.

[67] Poskosky, K. *Design of a Monocular Multi-Spectral Akin Detection, Melanin Estimation, and False Alarm Supression System.* Master's thesis, Air Force Institute of Technology, 2010.

[68] Ptitsyn, A. "Computational analysis of gene expression space associated with metastatic cancer." *BioMed Central Bioinformatics*, 10 Suppl 11:S6, 2009. URL `http://dx.doi.org/10.1186/1471-2105-10-S11-S6`.

[69] Ptitsyn, A., M. Hulver, W. Cefalu, D. York, and S.R. Smith. "Unsupervised clustering of gene expression data points at hypoxia as possible trigger for metabolic syndrome." *BioMed Central Genomics*, 7:318, 2006. URL `http://dx.doi.org/10.1186/1471-2164-7-318`.

[70] Rab, G., K. Petuskey, and A. Bagley. "A method for determination of upper extremity kinematics." *Gait Posture*, 15(2):113–119, Apr 2002.

[71] Reitberger, J., P. Krzystek, and U. Stilla. "Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees". *International Journal of Remote Sensing*, 29:1407–1431, March 2008. ISSN 0143-1161. URL `http://portal.acm.org/citation.cfm?id=1451047.1451058`.

[72] Robinette, K.M., S. Blackwell, H. Daanen, M. Boehmer, S. Fleming, T. Brill, D. Hoeferlin, and D. Burnsides. *Civilian American And European Surface Anthropometry Resource (CAESAR) Final Report, Volume I: Summary.* Technical report, United States Air Force Research Laboratory, 2002.

[73] Rössl, C., L. Kobbelt, H. Seidel, and I. Stadtwald. "Extraction of feature lines on triangulated surfaces using morphological operators". *In Proceedings of the AAAI Symposium on Smart Graphics*, 71–75. 2000.

[74] Roweis, S.T. and L.K. Saul. "Nonlinear dimensionality reduction by locally linear embedding". *SCIENCE*, 290:2323–2326, 2000.

[75] Royston, J. P. "An Extension of Shapiro and Wilk's W Test for Normality to Large Samples". *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 31(2):115–124, 1982.

[76] Royston, J.P. "Some Techniques for Assessing Multivarate Normality Based on the Shapiro- Wilk W". *Applied Statistics*, 32(2):121–133, 1983. URL `http://www.jstor.org/stable/2347291?origin=crossref`.

[77] Saul, L. K., K. Q. Weinberger, J. H. Ham, F. Sha, and D. D. Lee. "Spectral methods for dimensionality reduction". *Semisupervised Learning*, 1:293–308, 2006.

[78] Schünke, M., E. Schulte, L. M. Ross, U. Schumacher, E. D. Lamperti, J. Rude, M. Voll, T. C. Telger, and K. Wesker. *THIEME Atlas of Anatomy: General Anatomy and Musculoskeletal System.* THIEME, 2006.

[79] Shapiro, S.S. and M.B. Wilk. "An analysis of variance test for normality (complete samples)". *Biometrika*, 52(3-4):591–611, 1965.

[80] Sherbrooke, E.C., N.M. Patrikalakis, and E. Brisson. "An Algorithm for the Medial Axis Transform of 3D Polyhedral Solids". *IEEE Transactions on Visualization and Computer Graphics*, 2:44–61, March 1996. ISSN 1077-2626. URL `http://dx.doi.org/10.1109/2945.489386`.

[81] Shinagawa, Y. and T.L. Kunii. "Constructing a Reeb graph automatically from cross sections". *IEEE Comput. Graph. Appl.*, 11:44–51, November 1991. ISSN 0272-1716. URL `http://dx.doi.org/10.1109/38.103393`.

[82] Tenenbaum, J.B., V. Silva, and J.C. Langford. "A Global Geometric Framework for Nonlinear Dimensionality Reduction". *Science*, 290(5500):2319–2323, 2000.

[83] Trujillo-Ortiz, A., R. Hernandez-Walls, K. Barba-Rojo, and L. Cupul-Magana. "HZmvntest:Henze-Zirkler's Multivariate Normality Test". [WWW document] A MATLAB file, 2007. URL `URLhttp://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=17931`.

[84] Trujillo-Ortiz, A., R. Hernandez-Walls, K. Barba-Rojo, and L. Cupul-Magana. "Roystest:Royston's Multivariate Normality Test". WWW Document - MATLAB File, 2007. URL `http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=17811`.

[85] Uriarte, E.A. and F.D. Martn. "Topology Preservation in SOM". *World Academy of Science Engineering and Technology*, 15(October):52–55, 2006.

[86] Vacavant, A., D. Coeurjolly, and L. Tougne. "A framework for dynamic implicit curve approximation by an irregular discrete approach". *Graphical Models*, 71:113–124, May 2009. ISSN 1524-0703. URL `http://portal.acm.org/citation.cfm?id=1528928.1529069`.

[87] Vanco, M., B. Hamann, and G. Brunnett. "Surface Reconstruction from Unorganized Point Data with Quadrics". *Computer Graphics Forum*, 27:1593–1606, 2008.

[88] Wang, J. and J. Shan. "Segmentation of LIDAR Point Clouds for Building Extraction". *American Society for Photogrammetry Remote Sensing Annual Conference*. 2009.

[89] Werghi, N., Y. Xiao, and J. P. Siebert. "A functional-based segmentation of human body scans in arbitrary postures". *IEEE Transactions on Systems, Man and Cybernetics*, 36(1):153–165, 2006.

[90] Wu, G., F. Van der Helm, H. DirkJan Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A.R. Karduna, K. McQuade, X. Wang, F.W. Werner, and B. Buchholz. "ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion–Part II: shoulder, elbow, wrist and hand." *Journal of Biomechanics*, 38(5):981–992, May 2005.

[91] Wu, K. and M.D. Levine. "3D Part Segmentation Using Simulated Electrical Charge Distributions". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1223–1235, 1997.

[92] Xiao, Y., P. Siebert, and N. Werghi. "A Discrete Reeb Graph Approach for the Segmentation of Human Body Scans". *International Conference on 3-D Digital Imaging and Modeling*, 378–385. IEEE Computer Society, 2003. ISBN 0-7695-1991-1.

[93] Xiao, Y., P. Siebert, and N. Werghi. "Topological segmentation of discrete human body shapes in various postures based on geodesic distance". *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004)*, volume 3, 131–135. 2004.

[94] Zagoruiko, N. G. and G. S. Lbov. "Algorithms of Pattern Recognition in a Package of Applied Programs OTEKS". *4th International Joint Conference on Pattern Recognition*. 1978.

[95] Zagoruiko, N.G., V.N. Elkina, G.S. Lbov, and S.V. Emelianov. "Algorithms of Data Analysis in Application Package OTEKS". *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)*, 9:767–789, 1999.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
16–06–2011	Master's Thesis	Aug 2009 — Jun 2011

4. TITLE AND SUBTITLE

Estimating Anthropometric Marker Locations from 3-D LADAR Point Clouds

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Matthew J. Maier, Capt, USAF

5d. PROJECT NUMBER

10ENG300

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GE/ENG/11-27

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/RHPA (Julia Parakkat)
2800 Q Street, B824 WPAFB, USA 45433 937-255-0605,
Julia.Parakkat@wpafb.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RHPA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

An area of interest for improving the identification portion of the system is in extracting anthropometric markers from a Laser Detection and Ranging (LADAR) point cloud. Analyzing anthropometrics markers is a common means of studying how a human moves and has been shown to provide good results in determining certain demographic information about the subject. This research examines a marker extraction method utlizing principal component analysis (PCA), self-organizing maps (SOM), alpha hulls, and basic anthropometric knowledge. The performance of the extraction algorithm is tested by performing gender classification with the calculated markers.

15. SUBJECT TERMS

Anthropometric marker extraction, LADAR, SOM, CAESAR database, dismount characterization

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Michael J Mendenhall, Maj, USAF (ENG)
U	U	U	UU	113	19b. TELEPHONE NUMBER *(include area code)* (937)255-3636; Michael.Mendenhall@afit.edu